An ETF-based measure of Stock Price Fragility

Renato Lazo-Paz *

This version: December 27, 2023.

Abstract

A growing body of literature employs equity mutual fund flows to measure a stock's exposure to non-fundamental demand risk - stock price fragility. However, this approach may be biased by confounding fundamental information, potentially leading to underestimation of risk exposure. We propose an alternative method that incorporates readily available primary market data from exchange-traded funds (ETFs). This method significantly enhances the predictive power of fragility in forecasting stock return volatility. Moreover, our measure captures the influence of increased ETF activeness while partially capturing the effect of institutional investors' ownership on price return volatility. Additionally, our analysis reveals a decrease in the explanatory power of mutual fund-based fragility.

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 would like to thank David Brown for sharing their list of exchange-traded funds (ETFs). We are thankful for helpful feedback from Eser Arisoy, Markus Broman, Travis Box (discussant), Eliezer Fich, Agustin Hurtado, Maksim Isakin, Fabio Moneta, Maurizio Montone, Qianru Qi, Yinjie Shen, Andrea Vedolin, Shyam Venkatesan, Wei Wang, and seminar participants at the FMA 2023 Annual Conference, Cleveland State University, NEOMA Businness School, Nova School of Business and Economics (*PhD Pitch Perfect*), and The University of Ottawa.

1 Introduction

Classical asset pricing theories state that stock prices fluctuate because of fundamental shocks, such as news. This argument is based on the assumption that trading unrelated to a firm's fundamentals triggers a response by arbitrageurs who take the opposite side of the trade, canceling out any potential impact on security prices (e.g., Fama, 1965; Ross, 1976). However, extensive research has documented that trading driven by non-fundamental information (e.g., sentiment, noise, liquidity) can influence stock prices and that arbitrage activity faces various limitations that contribute to the persistence of mispricing.[1] While evidence shows that *non-fundamental demand shocks* influence asset prices, scholars continue to debate how to empirically measure a stock's exposure to these shocks.

Earlier research shows that stocks bought by mutual funds experiencing substantial inflows tend to underperform in the long run, whereas those sold by funds facing outflows tend to outperform (e.g., Coval and Stafford, 2007; Frazzini and Lamont, 2008). Moreover, Lou (2012) finds that price pressure resulting from mutual fund flow-driven trades contributes to the persistence of stock return momentum and mutual fund performance. This evidence has motivated a large body of literature to use investor flows to and from mutual funds as sources of exogenous non-fundamental price pressure.[2]

Building on this previous work, Greenwood and Thesmar (2011) developed the concept of *stock price fragility*. This measure combines information on an asset's ownership composition with data on the correlation between owners' non-fundamentally driven trades. These trades are proxied by mutual fund flows to

1 Seminal theoretical papers model 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 Wardlaw (2020) and Berger (2022) provide excellent recent discussions about the literature that relies on mutual fund flows as exogenous shocks to stock prices.

capture firm-level exposure to non-fundamental demand risk. Therefore, a stock is considered *fragile* if a few owners hold a large percentage stake (i.e., concentrated ownership) or if its owners face highly correlated non-fundamental demand shocks. This intuitive interpretation has prompted researchers to use this measure extensively.[3] Nonetheless, recent evidence has raised doubts about the empirical validity of mutual fund flows as instruments for non-fundamentally driven price pressure. Specifically, recent studies demonstrate that mutual fund flows motivate fund managers to perform discretionary trades[4] (Huang et al., 2022; Berger, 2022) and that such flows attract time-varying specialized demand from other mutual funds (Rzeznik and Weber, 2022).[5] Additionally, mutual fund managers actively hedge against the impact of common flows on fund size by tilting their portfolios toward low-flow-beta stocks, even at the expense of providing lower risk-adjusted returns (Dou et al., 2022). Theoretical models such as the influential model of Berk and Green (2004) argue that mutual fund flows reflect learning about mutual fund manager skills and thus do not necessarily reflect only non-fundamental demand. Overall, it is likely that the impact of mutual fund flows on prices cannot be exclusively attributed to non-fundamental demand. It also encompasses trades motivated by fundamental information.

The focus of this study is to provide an alternative method for estimating stock price fragility by employing data on exchange-traded funds (ETFs). Brown et al.

3 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), and equity issuance and repurchase activity (Massa et al., 2020). 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 that are driven by noise trading.

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) proportionally to the current weights of each asset in their portfolios.

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

(2021) introduce a model that links ETF primary market flows (i.e., the creation and redemption of ETF shares) to non-fundamental demand shocks.[6] The authors provide empirical evidence that supports their theoretical predictions. In light of this evidence, we propose an alternative method to estimate stock price fragility by employing ETF primary market flows and ownership composition data. This approach effectively overcomes many limitations associated with relying on mutual fund data while offering a more comprehensive scope by including a broader set of non-fundamental-driven sources of price variation. This is because ETFs are traded by a broad cross-section of market participants (i.e., retail traders, institutional investors, and Hedge funds), while mutual funds are mostly held by retail investors and households.

Our methodology provides three significant improvements over the existing method: 1) it relies on observable signals of non-fundamental demand not confounded by information about fund manager skills or fundamentally motivated trades (i.e., ETF flows); 2) it captures the impact of ownership and demand from both retail and institutional investors; and 3) it provides additional insights into the impact of the ETF industry's growth on asset prices.[7] Furthermore, in light of recent discussions concerning the impact of growing activeness in the ETF industry and the emergence of specialized attention-grabbing thematic ETFs (e.g., Easley et al., 2021; Ben-David et al., 2023), our study contributes by further exploring the effects of increased ETF activeness on asset prices.

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

7 Another potential advantage of an ETF-based fragility measure is that it can be estimated for a higher frequency (i.e., monthly), as opposed to the traditional mutual fund approach that relies on quarterly data. This approach could offer valuable insights into short-lived price dislocations, making it a promising avenue for future research.

Our analysis consists of two main parts. In the first part, we test the validity of our proposed methodology and compare it with the original estimation method. We begin by estimating the stock price fragility measure as in Greenwood and Thesmar (2011), G^{MF}, for the sample period used in that study (*in-sample*) and extend it until the last quarter of 2018 (*out-of-sample*). We then proceed to estimate the fragility measure employing only ETF data, G^{ETF}. Finally, in a regression setting, we test the ability of each measure to forecast future return volatility. In the second part of our analysis, we explore the factors that potentially make G^{ETF} superior measure and investigate the determinants of our prior findings. Specifically, we examine whether G^{ETF} captures the previously documented impact of institutional investors' ownership on volatility and whether increased ETF activeness helps explain our results.

We highlight four main empirical results. First, we find that the statistical and economic significance of G^{MF} in forecasting the next quarter's stock return volatility has significantly declined in the second part of our sample (2009-2018) - *out-of-sample*. Greenwood and Thesmar (2011) document that for the period between 1989 to 2008, an increase in G^{MF} fragility from 25th to the 75th percentile predicts an increase in daily volatility by 0.5%. Nevertheless, during the *out-of-sample* period, our estimation suggests that a comparable increase in fragility is associated with an expected rise in daily volatility of approximately 0.25%. While we do not focus on studying the determinants of this decline, we observe that this behavior coincides with a period during which the equity mutual fund industry has experienced significant outflows, as shown in Figure 1. Simultaneously, there has been substantial growth in the ETF industry in terms of trading volume and trading by a broader set of market participants (Dannhauser and Pontiff, 2019; Glosten et al., 2021; Easley et al., 2021). For instance, we estimate that by the last quarter of 2021, approximately 70% of Mutual funds and Investment advisors in the 13F institutional investors holding database included ETFs in their portfolios.

Second, we show that G^{ETF} strongly predicts the next quarter's stock return volatility in the later part of our sample period (2009 - 2018). Moreover, we find that when we include both G^{MF} and G^{ETF} in our regression model, only the coefficient of G^{ETF} remains positive and statistically significant. This evidence supports the conjecture that G^{ETF} provides information on fragility above and beyond that included in the G^{MF} measure. Our results align with evidence of an increase in ETF trading volume (Ben-David et al., 2017) and the integration of ETFs into both institutional and retail investors' portfolios (Dannhauser and Pontiff, 2019). Furthermore, our findings support the empirical evidence of Brown et al. (2021) and Davies (2022), indicating that ETF primary flows are indicators of non-fundamental demand shocks.

Third, we present evidence that G^{ETF} captures the influence of mid and small-sized institutional ownership on stock price volatility. In a recent study, Ben-David, Franzoni, Moussawi and Sedunov (2021) show that increased ownership by large- and mid-sized institutional investors predicts higher volatility and noise in stock prices. This effect arises from the granular nature of these institutions, where sub-units within large institutional investors tend to exhibit correlated trading behavior. This phenomenon, in turn, reduces the ability of institutional investors to diversify idiosyncratic demand shocks since correlated trades result in larger trading volumes, ultimately leading to more substantial price impacts We follow Ben-David, Franzoni, Moussawi and Sedunov (2021) specification and find that G^{ETF} remains statistically significant even when accounting for the impact of institutional investors' ownership on future stock price volatility. Furthermore, when G^{ETF} is incorporated into our regression analysis, the coefficient of mid-sized institutional ownership becomes statistically insignificant. We interpret this finding to be a consequence of the distinct ETFs ownership structure. Unlike mutual funds, which retail investors primarily own, ETFs are roughly equally owned by both retail and institutional investors (Dannhauser and Pontiff, 2019).[8] Moreover, we present additional evidence of the widespread adoption of ETFs by 13F institutional investors over time, particularly among investment advisors and transient institutions, that tend to have higher activity levels and shorter investment horizons. This fact can help explain why including G^{ETF} subsumes the explanatory power of mid-sized institutions.

Fourth, we document that the forecasting power of G^{ETF} on the next quarter's stock price volatility is mostly explained by *active* ETFs. It is possible that our results may be influenced by the comparison of two fundamentally distinct investment vehicles owing to their differing investment mandates. Equity mutual funds are actively managed, whereas ETFs were originally designed as passive vehicles with the primary objective of replicating a benchmark. We address this concern by estimating the activeness index of Easley et al. (2021) using our sample of ETFs. We corroborate the authors' findings in a broader sample of ETFs and show that ETFs have become, on average, more active in recent years.[9] Additionally, motivated by Easley et al. (2021) concerns that increased ETF activeness might negatively affect price informativeness by channeling active bets, we decompose the G^{ETF} into active and non-active components following the methodology outlined by Greenwood and Thesmar (2011). Our findings indicate that our results primarily stem from the active ETFs component. These results are consistent with Ben-David, Franzoni,

8 On appendix 6, we offer a comprehensive description of the incorporation of ETFs into the portfolios of 13F institutional investors. Our findings reveal a steady increase in the inclusion of ETFs, including leveraged/inverse ETFs, in institutional investors' holdings in recent years.

9 Easley et al. (2021) defines ETF activeness as being either in *form* or in *function*. A fund is active in *form* if it is designated to deliver out-performance or alpha. In *function* suggests that whether a fund is passively or actively managed, it can serve as a foundational component of an actively managed portfolio.

Moussawi and Sedunov (2021), who demonstrate that the expansion of the ETF industry has given rise to a multitude of specialized ETFs designed to cater to investors' extrapolation beliefs and prevailing investment trends. This phenomenon has led investors to allocate their wealth to already overvalued underlying stocks, exacerbating mispricing. When this mispricing is eventually corrected, it results in negative alphas for investors. Importantly, this evidence indicates that G^{ETF} measure can capture recent trends in the ETF industry that influence a stock's exposure to non-fundamental demand—an aspect largely overlooked by the G^{MF} measure.

Overall, our results are consistent with the argument that ETF primary markets flows provide valid signals of non-fundamental demand shocks (Brown et al., 2021) and that not only retail ownership but institutional investors' ownership contribute to stock return volatility (Bushee and Noe, 2000; Koijen and Yogo, 2019; Ben-David, Franzoni, Moussawi and Sedunov, 2021). Recent developments in the asset management industry, such as the rise of passive investing, increased accessibility to broader datasets, and advancements in theoretical frameworks and empirical evidence, call for a reevaluation of stock price fragility estimation. In this study, we address these developments and propose a revised fragility estimation method.

Our paper contributes to the ongoing discussion on the validity of mutual fund flows as instruments for non-fundamentally driven price variations. Specifically, we add to the growing literature that uses ETFs as a laboratory to study non-fundamental demand. Recent research has cast doubt on the empirical validity of two widely used approaches that rely on mutual fund data: one involving extreme outflows and the other employing a normalized measure, MFFLOWS. These approaches have been found fail to satisfy the conditions necessary to be considered valid instruments, and are not entirely orthogonal to fundamentals. Regarding the first approach, Huang et al. (2022) document that fire sales (i.e., those in which fund managers are forced to sell part of their holdings because of large outflows) contain fundamental information.[10] In essence, fund managers actively select which assets to sell rather than mechanically reducing all their positions, as previously assumed. Rzeznik and Weber (2022) document that the impact of fire sales on stock prices is negligible when mutual funds that hold the same stocks receive inflows. This suggests that specialized demand from these other funds mitigates the negative effects of fire sales by counteracting and purchasing these stocks. This evidence implies that the effects of fire sales are observable only in the absence of specialized demand. This conditional effect limits the suitability of fire sales as an adequate measure for capturing non-fundamental demand shocks In an influential paper, Wardlaw (2020) demonstrates that a widely used measure, the MFFLOWS of Edmans et al. (2012), is a direct function of realized returns during the outflow quarter. Moreover, the author shows that several documented results no longer hold once MFFLOW is corrected for this mechanical relationship. Similarly, Berger (2022) shows that the assumption that managers sell firms in proportion to portfolio weights induces selection bias in studies that employ the MFFLOW measure.[11] That is, it misallocates large price impacts to poorly performing illiquid firms with lower growth, which are, in fact, firms that fund managers avoid selling. Thus showing that the assumption does not hold true. Our study adds to this discussion by providing evidence that is in line with Brown et al. (2021) and Davies (2022). Specifically, our results reveal

10 Huang et al. (2022), show that when faced with large outflows, fund managers decrease only 43.9% of their holdings, and 37.4% of their positions remain unchanged. Surprisingly, the authors find that following large outflows, fund managers expand their holdings in 18.7% of securities, and such buys are more likely related to fundamentals since they can forecast future positive returns.

11 This refers to the *proportional trading assumption*. 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).

that when included in the estimation of price fragility, ETF primary market flows exhibit properties and outcomes consistent with reliable proxies for non-fundamental demand shocks.

In addition, our study contributes to the literature by examining the impact of non-fundamental demand on asset prices. Our research shows that an ETF-based measure of stock price fragility overcomes the limitations associated with using mutual fund data and effectively captures the influence of a broader range of investors on stock price volatility. Furthermore, our work adds to the growing body of literature investigating the effects of ETF activity on the volatility of their underlying assets (Ben-David et al., 2018) and the consequences of increased ETF activeness and heterogeneity of ETF products on stock prices (Easley et al., 2021; Davies, 2022; Ben-David et al., 2023). While extensive evidence exists on how ETFs can amplify the volatility of underlying stocks, our analysis extends these findings by considering ownership structure as a complementary factor. Our findings align with the insights of Israeli et al. (2017) regarding uninformed traders and ETFs, and Davies (2022) regarding to the role of ETFs, especially leveraged ETFs, in channeling investor gambling behavior. Our measure effectively captures these effects, which are often overlooked when relying solely on mutual fund data.

Recent studies emphasize the role of investor demand in explaining asset return patterns. In a pioneering work, Koijen and Yogo (2019) studied the impact of institutional investors and household ownership in determining stock demand elasticity and associated stock price volatility. Their findings indicate that while large institutional investors account for a substantial portion of market capitalization, mid- and small-sized institutional investors, as well as households, significantly contribute to stock price volatility. We believe that our measure contributes to this discussion by showing that an ETF-based fragility measure potentially captures the joint effect of retail and institutional stock ownership and demand shocks, channeled through ETF trading, on stock volatility. In this context, our results contribute to the current literature by revealing the role that institutional investor demand plays in non-fundamental demand shocks that ultimately influence stock prices.

The remainder of this paper is organized as follows. Section 2 describes the conceptual framework supporting our empirical approach. Section 3 describes the mutual funds and ETF data sources. Section 4 presents our main empirical results. Section 5 concludes the study and briefly discusses the implications of our results.

2 Conceptual framework

This section outlines the theoretical framework that motivates our empirical methodology. First, we review the literature that relates ETF primary market flows to non-fundamental demand shocks. Second, we provide an overview of recent studies that revisit the relationship between firms' ownership structure and non-fundamental demand risk, drawing links to our proposed methodology. Finally, we describe recent studies that discussed the limitations of mutual fund flows as a proxy of non-fundamental demand shocks and explain how an estimation of stock price fragility based on ETF data could effectively address and mitigate these concerns.

2.1 Non-fundamental demand shocks

Non-fundamental demand shocks cause market participants to trade an asset without regard to fundamental information about changes in future growth prospects or risk factors. Although the classic asset pricing theory regards these trades as *noise*, they can lead to deviations in asset prices from their intrinsic or fundamental values (De Long et al., 1990). The financial economics literature that investigates the factors behind such trades is extensive, and it can be broadly categorized into two main groups: noise/liquidity-driven (De Long et al., 1990; Wang, 1994) and sentiment-driven (Baker and Wurgler, 2006).[12] While the influence of non-fundamental demand shocks on asset prices has been extensively explored, identifying shocks orthogonal to any fundamental information remains an empirical challenge because fundamental values are unobservable. Many empirical studies have traditionally used mutual fund flows as a proxy for non-fundamental shocks. Nevertheless, this approach relies on assumptions that have faced scrutiny in recent years (e.g., Berger, 2022; Huang et al., 2022). To understand why ETF primary market flows offer clear and distinct signals of non-fundamental demand shocks, we briefly describe the redemption/creation mechanism underlying ETF trading. We then describe the link between this mechanism and the key insights from Brown et al. (2021) model.

ETFs are regarded as one of the most significant innovations in the asset management industry (Madhavan, 2014; Huang et al., 2020). Their remarkable success is commonly attributed to their cost efficiency and intraday liquidity.[13] However, a less recognized driver behind the rapid growth of the ETF industry is its superior tax efficiency compared to mutual funds, primarily because of the advantage of lower capital gain taxes (Moussawi et al., 2020).[14] These advantages led to the explosive growth of the ETF industry, resulting in the creation of a diverse range of investing products that track a wide array of benchmarks. This development provides investors with opportunities to gain exposure to both the broad market and specific

12 The literature on investor sentiment encompasses explanations grounded in concepts of both overreaction and underreaction (Barberis et al., 1998), gambling-like behavior (Kumar and Lee, 2006), and the disposition effect (Barber and Odean, 2000), among various other phenomena explored by the behavioral finance approach.

13 As of 2021, the US ETF market comprised around 2,570 funds, collectively accounting for a total of $7.2 trillion in net assets. In a global context, the total value of the worldwide ETF market reached $10.1 trillion (ICI, 2022 - available at https://www.icifactbook.org/pdf/2022_factbook.pdf)

14 Moussawi et al. (2020) document that the tax efficiency of ETFs relative to mutual funds increases long-term investors' after-tax returns by an average of 0.92% per year.

sectors and themes (Ben-David et al., 2023). As a result of the tremendous growth of the ETF Industry, roughly 35% of U.S. equity trading volume is attributable to ETFs (Glosten et al., 2021).

Adding to this distinguishing feature of offering investors intraday liquidity is the redemption/creation mechanism, which sets ETFs apart from other investment vehicles. This mechanism ensures that ETF shares expand or contract based on investors' demand. Because of the interaction between ETF share supply and investor demand, ETF share values may diver from the Net Asset Value (NAV) of the underlying securities that compose the benchmark (i.e., ETF premium). When such disparities occur, a specialized group of investors, referred to as Authorized Participants (AP), engage in trading activities involving the purchase and sale of large blocks of ETF shares with the ETF sponsor. The trading activity of APs corrects any arbitrage opportunities, ensuring that ETF intraday prices closely approximate the NAV of the underlying portfolio. This process is known as the creation-redemption mechanism or ETF primary market.

The creation/redemption process of ETFs on the primary market indicates excess demand from investors. When there is an increased demand for ETF shares, Authorized Participants (APs) acquire a block of new ETF shares from the ETF sponsor. This transaction involves transferring the basket of underlying securities to the sponsor and subsequently selling the newly acquired ETF shares in the secondary market. Conversely, the opposite process occurs when excess demand for the underlying assets that comprise the ETF's benchmark surpasses the NAV of the ETF shares. Brown et al. (2021) argue that this temporary dislocation between the ETF's share value and the NAV 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., *ETF flows*) allow researchers to observe these non-fundamentally driven trades. The author's model shows that in equilibrium, ETF flows do not contain information about fundamental information shocks.[15] Instead, they are the product of net excess demand in either the ETF shares or the ETF underlying assets. In other words, ETF flows act as a proxy for the magnitude and direction of non-fundamental demand shocks.

Brown et al. (2021) corroborate the predictions of their theoretical model by empirically showing that ETF flows forecast future asset returns that later reverse, and that this effect is stronger among leveraged and high-activity ETFs (those with more active primary markets). More recently, Davies (2022) expands this model to estimate a market-level Speculation Sentiment Index that captures aggregate speculative trades channeled through the trading activity of leveraged ETFs. His results are consistent with speculation sentiment causing market-wide price distortions that later revert.

Brown et al. (2021) relate their model to the well-known Berk and Green (2004) model which states that mutual fund flows are indicative of investors' learning and adapting behaviors concerning a manager's skill. Nevertheless, a significant distinction arises, given that ETFs are passively managed vehicles. Consequently, ETF flows do not reflect investors' learning of managerial skills. Instead, they reflect the competitive dynamics among Authorized Participants (APs) who exploit any misalignment between the value of ETF shares and their underlying NAV. This distinction serves as one of the primary advantages of an ETF-based fragility measure compared to that derived from mutual fund data. ETF flows lack discretionary skill-revealing information and rely on signals from arbitrage trading.

15 This is because, even though both the demand for ETF shares and the demand for the underlying assets contain fundamental information, this particular component does not directly contribute to the relative mispricing observed when the ETF premium emerges.

Overall, motivated by recent theoretical and empirical evidence, we argue that the arbitrage mechanism that characterizes the ETF primary market provides two main benefits for fragility estimation: (i) *observable non-fundamental demand shocks*: The creation and redemption process of ETF shares in the primary market offer distinct signals of non-fundamental demand shocks, which can be observed in data that tracks the number of outstanding ETF shares; (ii) *We do not need to rely on assumptions regarding fund managers behavior*: the mechanical correction of the misalignments between ETF NAV and underlying assets alleviates concerns regarding fund managers' discretionary decisions that might introduce fundamental information in fund flows.

2.2 Ownership structure and non-fundamental risk

Stock price fragility measures a security's exposure to *shifts* in non-fundamental demand by capturing the joint influence of ownership composition and the variance-covariance matrix of non-fundamentally-driven trades (i.e., *flows*) of asset owners. Greenwood and Thesmar (2011) introduced this measure based on a model that represents changes in an investor's portfolio assets as a function of two key motivations: i) those attributable to active rebalancing and ii) those arising from flow-driven trading. Then, assuming a stable relationship between aggregate flow-driven trades, a security's returns can be modeled as a function of price pressure due to flow-driven trades and an error term that captures information about the security fundamentals. If flow-driven demand cancels out across owners, prices should reflect only fundamental information. However, if non-fundamental demand is not solved, it has the potential to exert temporary non-fundamental pressure on prices.

Under the assumption of orthogonality between flow-driven trades and fundamental information, Greenwood and Thesmar (2011) concluded that the two key determinants of a security's return variance due to non-fundamental demand are: i) a vector representing the weight of each investor in that security (i.e., *ownership concentration*) and ii) the conditional variance-covariance matrix of flows originating from security owners (i.e., n*on-fundamental demand shocks*).

More recently, Ben-David, Franzoni, Moussawi and Sedunov (2021) expanded the model of Greenwood and Thesmar (2011) to study the relationship between large institutions' ownership and asset prices, specifically, return volatility. In principle, demand by large institutional investors influences stock return behaviors whenever shocks to these agents' portfolios are not easily diversified across their constituent subunits, influencing aggregate market outcomes (Gabaix, 2011). In other words, if funds under the same investment management firm exhibit some level of correlation in their trading activities when faced with external shocks to their holdings, then these institutions are considered granular. Their capacity to internally diversify these shocks is limited, ultimately resulting in a more pronounced market impact of their trades. Ben-David, Franzoni, Moussawi and Sedunov (2021) developed a model that relates asymmetric information and risk-averse market makers, linking asset managers' behavior to price dynamics. In their model, the variation in stock prices is represented as a function of three components: i) systematic aggregate shocks driving institutional trades, ii) fundamental idiosyncratic shocks, and iii) the effect of the ownership structure. Their main finding suggests that increased ownership by large institutional investors predicts higher volatility and noise in stock prices. Moreover, the authors find that institutional ownership has an impact on return volatility that is different from that of ownership concentration.

Overall, the theoretical models and empirical evidence reveal that stock return volatility is influenced by two key factors: ownership concentration and ownership by institutional investors. These variables have distinct effects on market dynamics. It is important to note that ownership by institutional investors, which constitutes the second element has been largely overlooked in Greenwood and Thesmar (2011) stock price fragility measure. This is because mutual funds are primarily held by households, while ETFs are owned and traded by a combination of institutional and retail investors (Dannhauser and Pontiff, 2019).[16] We argue that an ETF-based fragility measure is able to partially capture both effects, given the characteristics of investor ownership of ETF shares being split between retail and institutional investors. Additionally, the body of literature on the growing use of ETFs by institutional investors is expanding rapidly, documenting their role as a tool for actively gaining exposure to specific sectors (Easley et al., 2021) and to hedge against industry-specific risks (Huang et al., 2020). Furthermore, arbitrageurs commonly use ETFs to circumvent short-sale constraints (Karmaziene and Sokolovski, 2022; Li and Zhu, 2022).

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

Estimating stock price fragility presents two empirical challenges: i) identifying a source of independent shocks to stock prices that are orthogonal to firm fundamentals and are fully observable, and ii) access to comprehensive data on the ownership structure of assets. The first challenge, theoretically the most relevant, has been extensively explored in the financial economics literature. Beginning with Coval and Stafford (2007), numerous studies employ flow pressure from mutual fund sales as a proxy for non-fundamental price shocks.[17] Among the reasons for using mutual fund data were initial evidence showing that mutual funds mechanically reduce

16 On Appendix 0A4 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 findings in the literature by showing the widespread use of ETFs by institutional investors.

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

their portfolio holds when faced with significant outflows (i.e., fire sales) and the well-known fact that the vast majority of mutual fund share owners are households that are typically considered less financially sophisticated.[18] Motivated by this evidence, Greenwood and Thesmar (2011) relied on mutual fund data to estimate stock fragility.[19] Although using mutual fund flows as a proxy for non-fundamentally driven demand shocks has been a traditional approach in several empirical studies, recent papers have raised concerns about the assumptions we rely on when employing such an instrument. More specifically, i) the proportional trading assumption and ii) the absence of discretionary trades.

For mutual fund flows to serve as an adequate instrument for exogenous price changes, non-fundamental demand shocks should be transmitted to all securities within the fund portfolio. Thus, mutual fund holdings should expand and contract their current positions in response to a demand shock, thereby influencing the prices of their underlying securities. Berger (2022) shows that mutual fund managers, when faced with large outflows, do not sell shares of their portfolio firms in proportion to their current portfolio weights, as assumed by the MFFLOW measure of Edmans et al. (2012). Thus, when empirically tested, the *proportional trading assumption* does not hold and leads to significantly biased inferences. Berger (2022) show that mutual fund managers *systematically* avoid selling poorly performing, illiquid firms with lower growth.

Closely tied to the proportional trading assumption is the assumption that

18 According to the 2022 Investment Company Institute (ICI) Fact Book, more than 89% of mutual fund assets in the US were held by households.

19 It's important to highlight that the fragility measure incorporates all mutual fund flows and does not depend on the most commonly used MFFLOW measure introduced by Edmans et al. (2012). MFFLOW aims to capture forced selling activity following large mutual fund outflows. However, Wardlaw (2020) points out that this measure induces a mechanical relation between the measure and raw returns. While Greenwood and Thesmar (2011) approach does not directly suffer from this limitation, concerns that mutual funds flow convey fundamental information remain.

mutual fund flows do not incorporate fundamental information from discretionary trades of fund managers. Huang et al. (2022) reveal that during fire sales[20], mutual fund managers use fundamental information to direct a portion of their sales toward stocks with limited growth prospects (i.e., stocks with high short interest) while opting to sell fewer shares in stocks expected to beat earnings expectations in the next quarter. In line with these findings, Rzeznik and Weber (2022) find evidence that the negative impact of mutual fund fire sales on stock prices is negligible when specialized demand from other funds meets fire sale pressure. In other words, when active mutual funds hold a high valuation of a specific stock affected by fire sales from other funds, they opt to purchase that stock, effectively mitigating the adverse impact of selling pressure.

Overall, recent empirical evidence documents that even when mutual fund managers face selling pressure from significant outflows, they employ discretionary trades as a strategic response. These trades limit and concentrate the adverse effects of such demand shocks. In this process, fund managers introduce a blend of both fundamental (i.e., discretionary trades) and non-fundamental (i.e., expected or mechanical trades) information into their subsequent trades, ultimately influencing stock prices.

Recent studies suggest that, at most, mutual fund flows are noisy indicators of non-fundamental demand shocks. Motivated by this evidence and the theoretical and empirical findings of Brown et al. (2021), we argue that an ETF-based stock price fragility (G^{ETF}) remains unaffected by the documented concerns associated with mutual fund flows because i) ETF primary flows act as reliable signals for non-fundamental demand shocks, and ii) the mechanical arbitrage processes inherent to the creation and redemption of ETF shares mitigate concerns about discretionary trades conveying fundamental information. Furthermore, as previously discussed,

20 In the mutual fund literature, a fire sale event is typically defined as occurring when the fund experiences a net outflow equal to 5% or more of its total net assets (TNA).

this measure can capture the influence of institutional demand on asset prices — a factor that is completely overlooked by the current methodology.

We follow Greenwood and Thesmar (2011) and propose a fragility measure that employs only information (i.e., fund flows and ownership composition) from the ETFs.

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

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

We further expand the expression in Equation (1) to explicitly differentiate between Active and Passive ETFs as detailed 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, as detailed in the following Equation.

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

While this decomposition narrows its focus to ETF ownership alone, it presents the advantage of assessing the influence of passive and active ETFs on the measure. Moreover, this specification enables us to empirically investigate the concerns raised by Easley et al. (2021) regarding the impact of the increased activeness of the ETFs on price discovery. This is a key aspect to consider, given that the evolution of the ETF industry has been marked by the introduction of a wide variety of heterogeneous products (Ben-David et al., 2023). We argue that our measure helps to shed light on these open questions regarding the impact of ETF trading activity on overall market efficiency. While G^{ETF} represents a potential significant improvement in the estimation of stock price fragility, we are aware that it still has some limitations. Specifically, we rely on the assumption of uncorrelated liquidity-driven trades from investors outside our sample.

3 Data and variable construction

we first estimate the original measure of Greenwood and Thesmar (2011) to assess whether an ETF-based fragility measure proves to be a better measure. To create the required database of mutual funds and ETFs, we collected and combined data from several sources, as discussed in detail in the following section.

3.1 Mutual funds data

Our sample consists of US mutual funds from 1989 to 2018. Furthermore, in several tests, we partition the sample period into two distinct periods: from 1989 to 2008 and 2009 to 2018. To determine the sample periods, we followed two criteria. First, we closely follow Greenwood and Thesmar (2011) and begin our sample period from the last quarter of 1989 to the last quarter of 2008. This allowed us to replicate their estimations (i.e., *in-sample results*). Second, although the first US-listed ETF, the SPDR, was launched in 1993, ETFs 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 the end of Greenwood and Thesmar (2011) sample period. Thus, to test the explanatory power of our proposed measure, we focus on the latter part of our sample, starting in 2009, which allows us to capture the increase in ETF activity and perform an *out-*

of-sample test of the original fragility measure in the context of the rise of passive investing.[21]

We collect fund returns and total net assets (TNA) from the Center for Research in Security Prices (CRSP) Mutual Fund Database, We then collect mutual funds' quarterly holdings data from the Thomson/Refinitiv Mutual Fund Database (*s12*). We merged both databases by using the MFLinks database. As commonly done in the 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 and Pavlova and Sikorskaya (2023) to create the mutual funds holdings database.[22] Mutual funds with less than $ 5 million dollars in total net assets were excluded.[23] 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 to or above NYSE market capitalization decile 5.[24]

21 For instance, Madhavan (2014) highlights that 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.

22 We describe the merging of holdings databases and selecting mutual funds process in detail in Section OA2. of the Appendix.

23 While Greenwood and Thesmar (2011) does not explicitly impose this filter, we follow Fama and French (2010) and include the 5 million in TNA to control for the effects of incubation bias Evans (2010).

24 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.

3.2 Exchange-traded funds (ETFs) data

To create our primary ETF database, we begin by reviewing the list of ETF identifiers from Brown et al. (2021).[25] We extend this database to include ETFs up to the last quarter of 2018. We combined this data with information from Bloomberg and CRSP. From Bloomberg, we obtain data on outstanding shares and funds' net asset value (NAV). When data were missing or incomplete, we supplemented 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 ETF data sample covers the period from 2000 to 2018. In total, our sample includes 1,096 distinct ETFs for which we have both holdings and price/return data.

We impose the same filters on stocks in the ETF holdings database as those used in the mutual funds' sample to ensure comparability. Specifically, we retain stocks with market capitalization falling within the 5th decile or above of the NYSE breakpoint size deciles.

3.3 Estimating Fragility

We estimate stock price fragility as detailed in Equation (1). The two main components of the fragility measure are the 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 (ETF) portfolio allocation to stock i relative to the fund's total net assets (net asset value), as described in the following expression:

25 We thank David Brown for providing us with this data.

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

where $n_{i,j,t}$ is the number of securities i held by mutual fund (ETF) j at time t, P_{it} is the price of security i, and $a_{j,t}$ is the total j mutual fund (ETF) total net assets (net asset value).

3.3.1 MF-based Fragility (G^{MF})

For our mutual fund sample, we calculate the *percentage flows* for each mutual fund i at the end of quarter t as follows:

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

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. Because we employ the dollar positions of each fund in each security in matrix W, we require the covariance matrix Ω_t to be expressed in dollar terms.[26] We follow Greenwood and Thesmar (2011) and rescale the Ω_t matrix by funds assets at time t to obtain an estimate $\hat{\Omega}_t$:

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

For each quarter t, we calculate $\hat{\Omega}_{j,t}$ using a five-year rolling window estimation starting from 1984:Q1. Finally, fragility is estimated as shown in the following equation:

$$G^{MF_{it}} = \left(\frac{1}{\theta_{i,t}}\right)^2 W_{i,t}^{MF}\Omega_t^{MF}W_{it}^{MF}, \tag{3}$$

26 Greenwood and Thesmar (2011) warns about using dollar units to construct the variance-covariance matrix of flows since it would induce heteroskedasticity.

3.3.2 ETF-based Fragility (G^{ETF})

The elements of matrix W are estimated in 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. Similar to the methodology applied for MF-based fragility, we estimate ETF flows as percentage changes. In the context of the ETF primary market, this involves calculating the change in shares outstanding for each ETF k at each time t,

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

As performed with the mutual fund data, we normalize the ETF fund flows covariance matrix $\Omega_{k,t}$ as follows:

$$\hat{\Omega}_k^{ETF} = diag(NAV_{k,t})\Omega_{k,t}diag(NAV_{k,t})$$

To ensure consistency with the *MF-based* fragility estimation process, we estimated $\hat{\Omega}_k$ using a five-year rolling window[27]. Before 2005, ETF holdings represented only a negligible percentage of a stock's outstanding shares (Da et al., 2020). Consequently, utilizing data from this period would likely result in imprecise values for our measure. To address this concern and ensure the reliability of our estimations, we start reporting ETF-based fragility values from 2009 onwards. This approach guarantees the inclusion of a more substantial dataset and helps mitigate the po-

27 This specification differs from the one used by Greenwood and Thesmar (2011). They calculate the $\Omega_{k,t}$ variance-covariance flow matrix at time t by including all data from 1989 to each quarter t. We adopt a methodology in line with Francis et al. (2021) and Huang et al. (2022) and employ a five-year rolling window to estimate $\hat{\Omega}_k$. This approach accounts for the time-varying nature of the flow variance-covariance matrix and ensures the inclusion of the most up-to-date information. Huang et al. (2022) shows that varying the rolling-window estimation to two, three, or five years has little effect on the results.

tential for noisy results. We estimate the *ETF-based* fragility (G^{ETF}) based on the specification as in Equation (1).

[Figure 1 Here]

Figure 1 depicts the total new cash flows to mutual funds and ETFs. In the early sample period, mutual funds mostly experienced inflows. However, beginning in 2006, mutual funds on aggregate experienced outflows, as shown in Panel A. In contrast, as shown in Panel B, ETFs experienced significant inflows over the years, especially in the later part of the sample period. Our results are consistent with those of Dannhauser and Pontiff (2019).

[Table 1 Here]

Table 1 presents the descriptive statistics of our Mutual Funds (Panel A) and ETF (Panel B) samples. In any given year, our sample includes more mutual funds (1,134) than ETFs (334). The average ETF is larger in terms of assets under management (AUM) and holds a larger number of stocks. This difference is most likely driven by the presence of very large ETFs.[28] Thus, the median fund size provides a more accurate picture, showing that the median mutual fund ($ 58 million) is slightly larger than the median ETF ($ 48 million) 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, as shown in Panel B of Table 1. As described by Greenwood and Thesmar (2011), for fragility to be a reliable forecaster of future volatility, a firm's ownership composition should

28 Easley et al. (2021) document that by 2020, the three largest ETFs were: Vanguard Total Stock Market Shares Index ETF (VTI), the iShares S&P 500 Index ETF, and the SPDR (SPY) with assets under management of $216.4, $253.4, $337.2 billions, respectively.

not be too volatile from one quarter to the next. We test this assumption by estimating the autocorrelation coefficient of the number of owners. This is the number of funds that own the same stock. Panel C of Table 1 shows that the one-quarter-autocorrelation coefficient for the number of mutual fund owners is 0.861, while for ETFs is 0.832. Moreover, we observe that the autocorrelation coefficient value stays above 0.70 for both samples up to a lag of four quarters. Our results for the mutual fund sample closely follow those reported by Greenwood and Thesmar (2011). Moreover, we provide evidence that the ownership structure is highly persistent *also* for our sample of ETFs. These results provide additional evidence in favor of the suitability of ETF data for estimating stock price fragility.[29]

Table 2 shows the descriptive statistics of the variables that compose the fragility measure as well as for the square root of MF-based and ETF-based fragility.[30] Panel A of Table 2 shows that the number of mutual funds and ETFs holding the same stocks increased over time, particularly in the ETF sample for the later part of our sample period. On average, stocks within the mutual fund sample are held by 50 funds, whereas in the ETF sample, this figure averages approximately 25 funds.

[Table 2 Here]

Panel B of Table 2 provides insights into the time-series variation of flow volatility, estimated as the standard deviation of percentage mutual fund (ETF) flows. The volatility of mutual fund flows exhibited an increase in the initial segment

29 This requirement on the *persistence of ownership* of sample firms can also means that if we observe a fund's ownership of stock i on quarter t, we require that same stock i to be part of the funds portfolio on quarter *t+1*. Thus in principle, this is less of a concern for ETFs since index-tracking ETFs hold most of the securities than compress the benchmark index. This should be a major concern also for active ETF since most of such funds deviate from their benchmarks by changing their weighting scheme rather than the selection of stocks to hold (Madan, 2010; Easley et al., 2021)

30 Greenwood and Thesmar (2011) prefer to use the square-root of fragility because it is proportional to variance. Moreover, the authors define fragility as the conditional expected variance of flow-driven net buys into a stock.

of our sample period from 1989 to 2009, which is consistent with the findings of Greenwood and Thesmar (2011). However, volatility shows a notable decline in the out-of-sample period from 2010 to 2018. Conversely, the volatility of ETF flows experienced a substantial increase over the entire sample period, particularly in the later period of 2014-2018. A potential explanation for this behavior is the *flow hedging* activity of active equity funds. Dou et al. (2022) find that active equity funds hedge against common flows by tilting their portfolios toward low-flow beta stocks.

Concerning the correlation between fund flows, Panel C of Table 2 shows a decrease in the mean values for the mutual funds and ETF sample. Nonetheless, after an initial decrease, the correlation among ETF flows remained fairly stable for 2009-2018. It is worth mentioning that both the bottom (p25) and top (p75) quintiles of flow correlation are considerably similar for both mutual fund flows and ETF flows over the full sample. Panel D of Table 2 summarizes the square root value of fragility. Notably, from 1989 to 2009, the mean value for mutual fund fragility exhibited a substantial increase, soaring from 0.039 to 0.143. However, in the later part of our sample period, this value declined averaging 0.102. In contrast, the mean $\sqrt{G}$ continued to rise steadily for the ETF sample.

A potential concern is that the estimated fragility values may be influenced by potential differences in the characteristics of the stocks included in each sample. We address this concern in Table A1 of the Online Appendix. In this analysis, we sorted stocks into five quintile portfolios based on their MF-based G^{MF} (Panel A) and ETF-based G^{ETF} (Panel B) for each quarter t. Subsequently, we calculated the time-series averages of the cross-sectional means for various stock-level characteristics. Our findings confirm several results of Greenwood and Thesmar (2011). We observe that fragility does not exhibit a monotonous correlation with the number of owners. This underscores the notion that fragility is contingent on both the composition of ownership and the correlation between owners' trading decisions. Surprisingly, while we confirm that smaller firms and growth stocks with lower B/M ratios exhibit higher MF fragility, we do not find the same pattern when examining the quintiles for ETF fragility. This aligns with the findings of Brown et al. (2021) who noted that ETF flows convey information distinct from mutual funds, as ETFs are utilized by a diverse cross-section of investors, including retail investors, institutional investors, and hedge funds, thereby reflecting a broader range of trading decisions. Additionally, in figure A4 of the Online Appendix, we confirm Greenwood and Thesmar (2011) findings and observe a clear and positive correlation between ETF fragility and subsequent stock price volatility.

4 Empirical Results

In this section, we present our primary analysis to validate the proposed ETF-based fragility as a measure of non-fundamental risk. To this end, we test whether the measure is useful for forecasting flow-induced trading volatility in a regression framework. Additionally, we expand our initial setting to incorporate the influence of institutional investors' ownership on stock price volatility and explore the implications related to the proposed ETF-based fragility measure. Finally, we consider the heterogeneity of the ETF industry and decompose ETF-based fragility to explore the role of active and passive ETFs in our earlier findings.

4.1 Fragility and stock return volatility

For fragility to be a useful measure of non-fundamental risk, it must forecast mutual fund (ETF) induced trading stock return volatility. We test this predictive power by estimating the following Fama and MacBeth (1973) regression.[31]

$$\sigma_{i,t+1} = \alpha + \beta\sqrt{G}_{i,t} + \delta Z_{i,t} + \mu_{i,t} \tag{4}$$

Equation (4) follows the main specification employed by Greenwood and Thesmar (2011), where $\sigma_{i,t+1}$ is the one-quarter-ahead standard deviation of daily stock returns. $Z_{i,t}$ 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 β measures the relationship between the current quarter's fragility and the next quarter's stock return volatility. Therefore, a positive and statistically significant value of β indicates that an increase in stock fragility in the current quarter would forecast an increase in stock return volatility in the next quarter. Table 3 presents the results from the regression model. We first predict future volatility using $\sqrt{G^{MF}}$ and its components across the entire sample period. The results of this test are presented in the first four columns of the table. To evaluate the *out-of-sample* performance of $\sqrt{G^{MF}}$, we replicate these four regression specifications for the latter portion of the sample period, spanning from 2009 to 2018. For comparability and to test our proposed ETF-based fragility measure, we run the same regression specifications on $\sqrt{G^{ETF}}$ for the same period.

[Table 3 Here]

Column (1) of Table 3 provides a first formal test of the relationship between fragility and future volatility for the sample period between 1989 and 2018. Consistent with previous findings, a positive and statistically significant relationship exists

31 We perform Fama and MacBeth (1973) regressions to control for the effect of common trends like increasing ownership of Mutual Funds and ETFs (Da et al., 2020).

between $\sqrt{G^{MF}}$ and next-quarter daily return volatility. Nonetheless, it's worth noting that the reported coefficient is considerably smaller than the value documented by Greenwood and Thesmar (2011) who reported a value of β equal to 0.696. In our analysis, we find this coefficient to be 0.459. Furthermore, when focusing on the latter part of our sample, as reported in Column (5), we observe a substantially reduced coefficient of 0.325, almost half of the coefficient reported by Greenwood and Thesmar (2011).

In columns (2) and (3), we examine the relationship between specific components of fragility, namely ownership (IO) and concentration, and expand the initial specification by introducing additional control variables. The results indicate a positive relationship between mutual fund ownership and future volatility[32], and that the explanatory power of fragility extends beyond pure ownership concentration, as proxied by the Herfindahl index. In column (4), we check whether the predictive power of $\sqrt{G^{MF}}$ remains robust when accounting for a comprehensive set of control variables, including the lagged dependent variable. This is important because volatility tends to exhibit a high persistence over time. The results reveal that the coefficient of fragility, denoted as β, decreases significantly to 0.072 (t-stat = 2.75). Moreover, if we focus on the latter part of our sample, the coefficient drops further to 0.018, reaching only marginal significance at the 10% level (t-stat = 1.70). These findings suggest that the forecasting power of $\sqrt{G^{MF}}$ on volatility significantly diminishes over time.

We repeat the analysis conducted in columns (1) to (4) using the ETF-based fragility and report our findings in columns (9) to (12). Our initial test shows that $\sqrt{G^{ETF}}$ is a strong positive predictor of the next-quarter standard deviation of daily stock returns with a β equal to 0.825 (t-stat = 7.76). Notably, this coefficient is

32 As previously documented by Sias (1996) and Bushee and Noe (2000).

significantly higher than that of $\sqrt{G^{MF}}$ for the same period, which stands at 0.325 (t-stat = 8.75). In column (10), we corroborate the findings of Ben-David et al. (2017) regarding the positive relationship between higher ETF ownership and increased volatility. Interestingly, even when we incorporate the full set of control variables, as shown in column (12), the relationship between $\sqrt{G^{ETF}}$ and future volatility remains strongly positive and statistically significant, as we obtain a coefficient value of 0.338 (t-stat = 5.93) Our results provide evidence that an ETF-based measure of fragility is useful in forecasting next quarter volatility. Moreover, our estimates indicate that the original measure of Greenwood and Thesmar (2011) lost forecasting power over time.

[Table 4 Here]

Next, we more directly investigate the conjecture that ETF-based fragility is a robust measure of non-fundamental demand risk. To this purpose, Table 4 presents an analysis of the volatility predictors for the later part of the sample period. We report the results of regressions in which we assess the influence of both $\sqrt{G^{MF}}$ and $\sqrt{G^{ETF}}$, along with a set of control variables, on the next-quarter daily return volatility. As previously mentioned, ETFs exhibit a distinct ownership composition to mutual funds, held nearly equally by households and institutional investors. Therefore, it can be anticipated that the effect of non-fundamentally driven demand captured by G^{ETF} differs from that of G^{MF}. In other words, while the ETF-based measure may capture a similar component to the MF-based fragility, namely, retail investors' demand, it is also possible that it incorporates the influence of institutional investors' demand.

To explore these differences, we repeat the analysis reported in Table 3 including both $\sqrt{G^{ETF}}$ and $\sqrt{G^{MF}}$ simultaneously. In Column (1) of Table 4, we test the joint effect of both fragility measures on future volatility. While we observe that daily volatility is positively and statistically significantly correlated with both measures, the coefficient of $\sqrt{G^{ETF}}$ is significantly larger than that of $\sqrt{G^{MF}}$. Moreover, the coefficient of $\sqrt{G^{MF}}$ is smaller than that reported in Column (5) of Table 3. This finding suggests that the ETF-based fragility measure captures, at least to some extent, an effect similar to but above and beyond that measured by the MF-based fragility. Column (2) shows that the mutual funds and ETF ownership are both positively correlated with future volatility. Columns (3) and (4), include control variables to test the robustness of our findings. The results show that, with the full suite of controls, the coefficient on $\sqrt{G^{MF}}$, 0.009, is no longer statistically different from zero (t-stat = 1.03). In contrast, the coefficient of $\sqrt{G^{ETF}}$, 0.426, remains highly significant (t-stat = 7.95). Overall, this evidence is also in line with Brown et al. (2021) findings that ETF flows include information distinct from the information in mutual fund flows.

[Table 5 Here]

To address concerns regarding the possibility that the regression settings influence our results, we also present panel fixed effects estimates, consistent with previous studies (Ben-David, Franzoni, Moussawi and Sedunov, 2021; Friberg et al., 2022). We include firm and year-quarter fixed effects and adjust the standard errors for clustering at the firm level. The sample period employed in this analysis spans from the first quarter of 2009 to the last quarter of 2018. We follow Friberg et al. (2022) specification and test the relationship between stock price fragility and future stock price fragility within three subsets: (i) the full sample, (ii) a subset comprising observations with a minimum of 20% institutional ownership, and (iii) a sample of firms with market capitalization above the median. These subsets are designed to assess the robustness of our findings, ensuring that they are not influenced or concentrated in firms with dispersed and relatively low levels of institutional ownership or by smaller firms. We report the results of this analysis in Table 5.

Our findings reported in columns (1), (5), and (9) closely match those of Friberg et al. (2022).[33] We observe that $\sqrt{G^{ETF}}$ is positive and statistically significant in all three subsets, as detailed in columns (2), (6), and (11). Moreover, in line with our previous findings, the magnitude of the coefficient of $\sqrt{G^{ETF}}$ is significantly larger than that of $\sqrt{G^{MF}}$. We then include two sets of control variables: those used by Friberg et al. (2022) (natural log of market capitalization and the inverse of stock price) and those employed by Greenwood and Thesmar (2011) as specified in Table 4. In columns (4), (8), and (12) we observe that when including the set of controls specified by Greenwood and Thesmar (2011) in the regressions, our results are similar to those obtained in the Fama-Macbeth regressions: $\sqrt{G^{MF}}$ loses all statistical significance while $\sqrt{G^{ETF}}$ remains positively and statistically significantly related to future stock price fragility. In summary, this analysis provides further evidence that the proposed ETF-based fragility measure is a robust and strong predictor of future return volatility.

Lou (2012) was among the first to propose a capital-flow-based explanation for some return predictability patterns. By aggregating flow-induced trading by mutual funds, the author finds that such demand shocks can partially explain stock price momentum. More recently, Li (2022) documents that price pressure from mutual fund investor demand explains approximately 30% of fluctuations in the Fama-French size and value factors. Thus, fragility may predict the volatility of risk factors

33 The disparities we observe might potentially be attributed to differences in sample periods since Friberg et al. (2022) conducted their analysis spanning from 2001 to 2017. In untabulated results, we replicated our analysis for this identical time frame and obtained results that closely mirror those reported.

themselves.[34] Thus, we also explore the relationship between fragility and volatility of returns in excess of several asset pricing factors. For excess return volatility, we estimate risk-adjusted returns using three models: (1) market-adjusted returns, (2) the Fama and French (1993) three factors model, and (3) the Fama and French (1993) model augmented with the Carhart (1997) momentum factor. Additionally, we estimate the DGTW-adjusted returns as in Daniel et al. (1997). The results of this analysis are shown in Table 6.

[Table 6 Here]

In Panel A, our results corroborate those of Greenwood and Thesmar (2011) as we observe that the coefficient of $\sqrt{G^{MF}}$ F is slightly smaller than that obtained when analyzing total return volatility. Furthermore, we note a significant decrease in the magnitude of this coefficient in the latter part of our sample period, 2009-2018. We observe a similar pattern in the ETF sample, as shown in the first part of Panel B. Subsequently, we examine the relationship when we include both fragility measures simultaneously. We see that the coefficients of $\sqrt{G^{ETF}}$ are significantly higher than those of $\sqrt{G^{MF}}$. Moreover, the inclusion of $\sqrt{G^{MF}}$ only marginally reduced the coefficient of $\sqrt{G^{ETF}}$. These results highlight the statistically significant association between fragility and excess return volatility for both fragility measures. However, it is worth noting that the ETF-based measure exhibits a stronger predictive power.

In summary, the analysis carried out in this section is consistent with the argument that an ETF-based fragility measure strongly predicts future stock return volatility. These empirical observations are consistent with the evidence that shows that ETF primary market flows reflect non-fundamental demand shocks.

34 Following this rationale, Greenwood and Thesmar (2011) argue that the predictability of fragility on *excess return volatility* is expected to yield weaker results.

4.2 Fragility and institutional investors' ownership

In this section, we study the determinants of the superior forecasting power of G^{ETF} on stock return volatility. We argue that ETF primary market flows potentially channel excess demand not mainly from one class of investors, as is the case with mutual fund flows, but from a broader cross-section of market participants including retail investors, institutional investors, and hedge funds.

[Figure 2 Here]

Figure 2 shows a significant increase in 13F institutional investors over the years, paralleled by a corresponding rise in the adoption of ETFs in their portfolios. By the end of 2000, approximately 20% of institutional investors included at least one ETF in their holdings. However, by the end of 2020, this proportion had surged to approximately 70% of all institutional investors. Interestingly, a similar exponential growth is observable in the case of leverage and inverse-leverage ETFs.[35]

[Figure 3 Here]

In Figure 3, we expand this evidence and analyze the time-series adoption of ETFs in 13F institutional investors' holdings. Moreover, we break down 13F data into different investor types and observe that investment advisors, mutual funds, quase-indexers, and transient institutions are among the institutional investors that have the most extensively incorporated ETFs in their portfolios. Interestingly, both short- and long-horizon investors, as defined by Yan and Zhang (2009), have similarly integrated ETFs into their portfolios, with a recent trend of heightened adoption

35 Similarly to traditional ETFs, Leverage ETFs offer exposure to a wide set of benchmarks. However, their replication method includes using derivatives. This mechanism allows ETF 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 offer negative exposure by obtaining a negative multiplier of the benchmark return, for instance -1.5x -2x of the return.

by long-term investors. In Appendix 6, we present data on leveraged and inverse-leveraged ETFs. We document that long-term investors, investment advisors, and quasi-indexers have consistently incorporated this investment vehicle into their portfolios over the years. By the end of 2021, nearly 50% of investors in each of these categories reported having at least one leverage or inverse-leveraged ETF in their portfolios. In summary, our findings confirm the findings in the literature by highlighting the widespread inclusion of ETFs in 13F institutional investors' holdings.

In a recent study, Ben-David, Franzoni, Moussawi and Sedunov (2021) shows that increased stock ownership by large institutional investors induces higher return volatility and greater noise in stock prices. This heightened volatility is primarily attributed to investors' inability to diversify idiosyncratic shocks among their subunits. In other words, subunits within large institutional investors tend to exhibit correlated behavior when faced with such shocks, amplifying their impact on asset price volatility. Considering the widespread inclusion of ETFs in the portfolios of 13F institutional investors, it is plausible that an ETF-based fragility measure may capture the price pressure resulting from institutional investors trading ETFs.

Prior literature shows that institutional investors engage in ETF trading for diverse reasons. For instance, Huang et al. (2020) show that hedge funds regularly implement a *long-the-stock/short-the-ETF* strategy relying on industry ETF to hedge their industry risk exposure. Similarly, Karmaziene and Sokolovski (2022) and Li and Zhu (2022) find evidence that arbitrageurs employ ETFs to circumvent short-sale bans and constraints. On the contrary, Sherrill et al. (2017) document a negative association between large ETF positions and mutual fund performance. The authors find that underperformance is mostly due to mutual funds' poor timing ability to implement investment strategies based on ETFs. Sherrill et al. (2020) show that many active mutual funds hold passive ETFs to reduce their cash holdings while relying on active ETFs to enhance fund performance. Nevertheless, the evidence supporting the latter proposition is somewhat limited.

In this section, we test the hypothesis that an ETF-based measure of stock price fragility effectively captures the impact of institutional investor trading. We introduce variables within a regression framework that considers ownership by large-, mid-, and small-sized institutional investors based on their assets under management. Our objective is to assess the influence of these variables on the predictive power of G^{MF} and G^{ETF} on future return volatility. We follow Ben-David et al. (2023) specification and perform the following panel regression:

$$\sigma_{i,t+1} = \beta_1 \mathrm{TopIO}_{i,t} + \beta_2 \mathrm{MidIO}_{i,t} + \beta_3 \mathrm{BottomIO}_{i,t} + \delta Z_{i,t} + \beta_4 G_{i,t} + \alpha_i + \theta_t + \mu_{i,t} \quad (5)$$

where $\sigma_{i,t+1}$ is the next quarter t stock i volatility. $\mathrm{TopIO}_{i,t}$ 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. $\mathrm{BottomIO}_{i,t}$ represents the aggregate stock's i ownership of the smallest institutional investors whose aggregate money holdings value equals that of the top institutions. $\mathrm{MidIO}_{i,t}$ is collective ownership by 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). α_i is the stock fixed effect, and θ_t is the time (calendar year-quarter) fixed effect.

[Table 7 Here]

Table 7 shows the results for two specifications: considering the Top 3 and Top 10 institutional investors. *Top IO* represents the aggregate ownership of the largest institutional investors in a given stock. For the *top 3 Institutions* specification, we sum the ownership of the top 3 institutions, whereas for the *top 10 Institutions* we sum the ownership of the top 10 institutions. We also perform the regression for the full sample and repeat the analysis for the later part, 2009 - 2018. In columns (1) and (1), our results closely follow those reported by Ben-David, Franzoni, Moussawi and Sedunov (2021). We observe a positive and statistically significant association between ownership by large and medium-sized institutional investors and stock volatility. This relationship is negative for bottom institutional ownership, consistent with the view that large investors affect volatility. Additionally, the coefficient on G^{MF} is also positive and significantly related to future stock price volatility.

Ben-David, Franzoni, Moussawi and Sedunov (2021) argue that including stock price fragility has little impact on their analysis because each measure captures two partially independent effects. In other words, the influence of concentration (i.e., fragility) and large institutional investors' limitations in diversifying away demand shocks to their holdings (i.e., granularity) have different impacts on stock price volatility. As previously discussed, we argue that an ETF-based fragility measure partially channels the effect of Institutional Ownership on stock price volatility given that ETFs are owned and traded by both retail and institutional investors. In Column (4), we test this hypothesis and replace G^{MF} with G^{ETF}. For comparability, we limit our analysis to the second part of our sample (2009-2018). We find that the coefficient on large institutional ownership and G^{ETF} are positive and statistically significant. However, the coefficients of the *mid* and *bottom IO* are smaller and indistinguishable from zero. This effect is observable if we change our setting and observe the top 10 institutional ownership, as detailed in Column (7).

[Table 8 Here]

In columns (5) and (8) we add both to our main regression model and find results similar to those documented previously. That is, the coefficient of G^{MF} loses statistical significance, while G^{ETF} remains economically and statistically significant. We replicate the results for the alternative grouping of top institutional investors, specifically the Top 5 and Top 7 Institutions in Table 8. Our results remain qualitatively the same.

Our evidence confirms our assumption that G^{ETF} measure partially captures the effect of institutional investors' demand on volatility, which G^{MF} does not consider. This analysis suggests that overlooking the impact of institutional ownership on return volatility could introduce a significant bias in the estimation of stock price fragility.

4.3 ETF activeness and stock price volatility

A valid concern in our empirical analysis is that we combine data from two *distinct* investment vehicles. In terms of their investment mandates, we compare *active* mutual funds while ETFs are, in principle, passively managed. We follow Easley et al. (2021) and estimate their *Activeness Index* to determine which fraction of our sample of ETFs can be considered active.

$$\mathrm{ActivenessIndex_{i,t}} = \sum_{s=1}^{N} w_{i,s,t} - w_{market,s,t} \tag{6}$$

We illustrate the composition of ETFs and their trading activity based on the median level of their *activeness index*. ETFs with activeness index values above the median value are considered active. Those below the median value are classified as passive. ETFs with values located in the top (bottom) quintile are considered very active (passive). In Table 9 we report descriptive statistics on the activeness index value for our full sample, for several subsets, and grouped by number of funds and aggregate assets under management (AUM).

[Table 9 Here]

The mean activeness index value in our sample is between 87% to 90%. In other words, the mean ETF in our sample is highly active. If we focus on the number of funds, over 94% are classified as either moderately active or very active. In terms of AUM, approximately 70% is managed by active ETFs. For our larger sample of ETFs, we corroborate the findings of Easley et al. (2021) and show that most ETFs can be classified as active investment vehicles. Furthermore, our findings align with those of Ben-David, Franzoni, Kim and Moussawi (2021), who documented that the evolution of the ETF industry has been marked by the emergence of niche, highly specialized products, including sector, thematic, industry, and smart-beta ETFs.

Is it possible that more active ETFs drive our results?. While Brown et al. (2021) do not distinguish between ETFs based on their activeness index value, Easley et al. (2021) expressed concerns about the potential negative impact of the increasing activeness of ETFs on price discovery. Thus, it is plausible to consider that more active ETFs could play a particularly significant role in propagating fragility, as they may attract a greater number of short-term, speculative trades. We follow Easley et al. (2021) and split our sample according to the activeness index (50% threshold), since this cutoff most likely includes both *active-in-form* and *active-in-function* ETFs. Then, we re-estimate the ETF fragility as detailed in Equation 2 and replicate the main specifications of Tables 4 and 6, considering the decomposition of G^{ETF} into active and passive parts.

[Table 10 Here]

Column (1) of Table 10 shows that most of the observed relationship between G^{ETF} and volatility stems from the 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 raised by Easley et al. (2021) regarding the role of increased ETF activeness in price informativenes. We show that the active ETF component of the ETF-based fragility measure is responsible for most of the observed relationship between fragility and future stock return volatility.

5 Conclusion

Wardlaw (2020) advocates reevaluating the empirical approach employed to measure non-fundamental price changes. The author raises concerns about the use of noisy, low-frequency data, such as mutual fund flows, for this purpose. Moreover, recent evidence challenges assumptions supporting the use of mutual fund flows. Huang et al. (2022) demonstrates that mutual fund managers exert discretionary trades that convey fundamental information during fire sales. Similarly, Berger (2022) tests the assumption that when faced with large outflows, mutual fund managers sell firms in proportion to portfolio weights; thus, no ability or skill from fund managers is included, and finds that this does not hold when empirically tested. More importantly, the author shows that relying on those assumptions can significantly affect the regression results, questioning inferences drawn from such analyses.

In our study, we turn to ETF primary market flows. Motivated by empirical and theoretical evidence showing that they clearly signal non-fundamentally driven demand shocks (Brown et al., 2021), we document that relying on this data significantly improves the estimation of stock price fragility (Greenwood and Thesmar, 2011) while avoiding the criticism and limitations surrounding the use of mutual fund flows data. We find that an ETF-based fragility measure strongly predicts price volatility. Moreover, given the growing concentration of equity holdings in a select few institutional investors, whose ownership ties closely to stock return volatility as evidenced in the literature (Koijen and Yogo, 2019; Ben-David, Franzoni, Moussawi and Sedunov, 2021), it is plausible that a fragility measure based on mutual funds may fail to account for this effect. Our findings show that an ETF-based fragility measure partially captures the effects of institutional ownership on stock volatility. These results are supported by evidence of increased ownership of ETFs by institutional investors (Dannhauser and Pontiff, 2019). Additionally, we address recent concerns about the effect of increased ETF activeness on return's volatility (Easley et al., 2021) and show that most documented effects stem from active ETFs. Overall, our findings offer a comprehensive view of the underlying dynamics of stock price fragility.

Our results have implications for empirical asset pricing studies. In particular, they inform the debate on the impact of non-fundamentally driven demand shocks on stock return volatility. Although our approach does not completely resolve the limitations associated with empirically estimating stock fragility, it represents a method not affected by many criticisms surrounding the use of mutual fund flows, thus offering researchers a more accurate proxy for assessing firm-level exposure to non-fundamental demand risk. In the broader context, our findings contribute to the ongoing discussion within the literature examining the repercussions of the rise in passive investments on overall market efficiency, a matter of great interest to policymakers and investment managers.

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.

Barber, Brad M. and Terrance Odean (2000), 'Trading is hazardous to your wealth: The common stock investment performance of individual investors', *Journal of Finance* **55**(2), 773–806.

Barberis, Nicholas, Andrei Shleifer and Robert Vishny (1998), 'A model of investor sentiment', *Journal of Financial Economics* **49**(3), 307–343.

Ben-David, Itzhak, Francesco A. Franzoni, Byungwook Kim and Rabih Moussawi (2021), 'Competition for Attention in the ETF Space', *SSRN Electronic Journal* **3**, 103–111.

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.

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.

Da, Zhi, Xing Huang and Lawrence J. Jin (2020), 'Extrapolative beliefs in the cross-section: What can we learn from the crowds?', *Journal of Financial Economics* .

Daniel, Kent D., David Hirshleifer and Avanidhar Subrahmanyam (2001), 'Overconfidence, arbitrage, and equilibrium asset pricing', *Journal of Finance* **56**(3), 921–965.

Daniel, Kent, Mark Grinblatt, Sheridan Titman and Russ Wermers (1997), 'Measuring Mutual Fund Performance with Characteristic-Based Benchmarks', *The Journal of Finance* **52**(3), 1035–1058.

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

Davies, Shaun William (2022), 'Speculation Sentiment', *Journal of Financial and Quantitative Analysis* **57**(7), 2485–2515.

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.

Dou, Winston Wei, Leonid Kogan and Wei Wu (2022), 'Common Fund Flows: Flow Hedging and Factor Pricing', *SSRN Electronic Journal* .

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.

Evans, Richard B. (2010), 'Mutual fund incubation', *Journal of Finance* **65**(4), 1581–1611.

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 (2010), 'Luck versus Skill in the cross-section of mutual fund returns', *Journal of Finance* **65**(5), 1915–1947.

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.

Frazzini, Andrea and Owen A Lamont (2008), 'Dumb money: Mutual fund flows and the cross-section of stock returns$', *Journal of Financial Economics* p. 24.

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* .

Gabaix, Xavier (2011), 'The Granular Origins of Aggregate Fluctuations', *Econometrica* **79**(3), 733–772.

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 '.

Israeli, Doron, Charles M.C. Lee and Suhas A. Sridharan (2017), *Is there a dark side to exchange traded funds? An information perspective*, Vol. 22, Review of Accounting Studies.

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.

Kumar, Alok and Charles M.C. Lee (2006), 'Retail investor sentiment and return comovements', *Journal of Finance* **61**(5), 2451–2486.

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.

Madan, Dilip B. (2010), 'Stochastic Processes in Finance', *Annual Review of Financial Economics* **2**(1), 277–314.

Madhavan, Ananth (2014), 'Exchange-Traded Funds: An Overview of Institutions, Trading, and Impacts', *Annual Review of Financial Economics* **6**(1), 311–341.

Massa, Massimo, David Schumacher and Yan Wang (2020), 'Repurchases for Price Impact: Evidence from Fragile Stocks', *SSRN Electronic Journal* (435).

Merton, Robert C (1971), 'Optimum consumption and portfolio rules in a continuous-time model', *Journal of Economic Theory* **3**(4), 373–413.

Moussawi, Rabih, Ke Shen and Raisa Velthuis (2020), 'The Role of Taxes in the Rise of ETFs', *SSRN Electronic Journal* (September).

Pavlova, Anna and Taisiya Sikorskaya (2023), 'Benchmarking Intensity', *Review of Financial Studies* **36**(3), 859–903.

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', *SSRN Electronic Journal* .

Sherrill, D. Eli, Sara E. Shirley and Jeffrey R. Stark (2017), 'Actively managed mutual funds holding passive investments: What do ETF positions tell us about mutual fund ability?', *Journal of Banking and Finance* **76**, 48–64.

Sherrill, D. Eli, Sara E. Shirley and Jeffrey R. Stark (2020), 'ETF use among actively managed mutual fund portfolios', *Journal of Financial Markets* **51**(xxxx), 100529.

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.

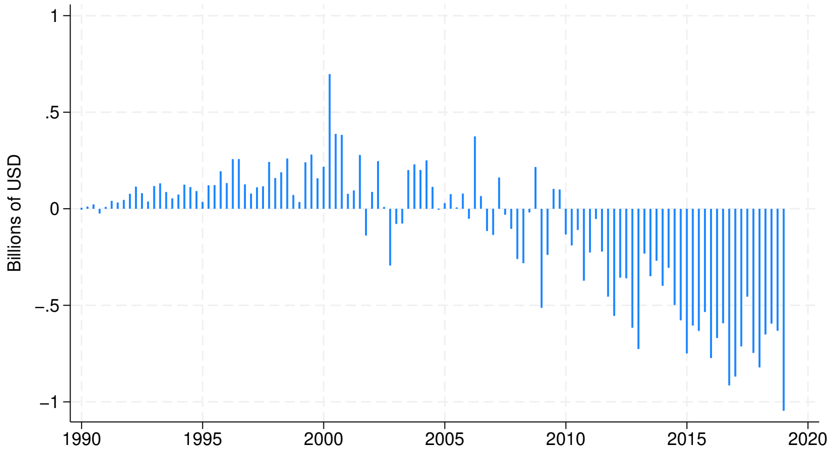
Wardlaw, Malcolm (2020), 'Measuring Mutual Fund Flow Pressure as Shock to Stock Returns', *The Journal of Finance* **75**(6), 3221–3243.

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.

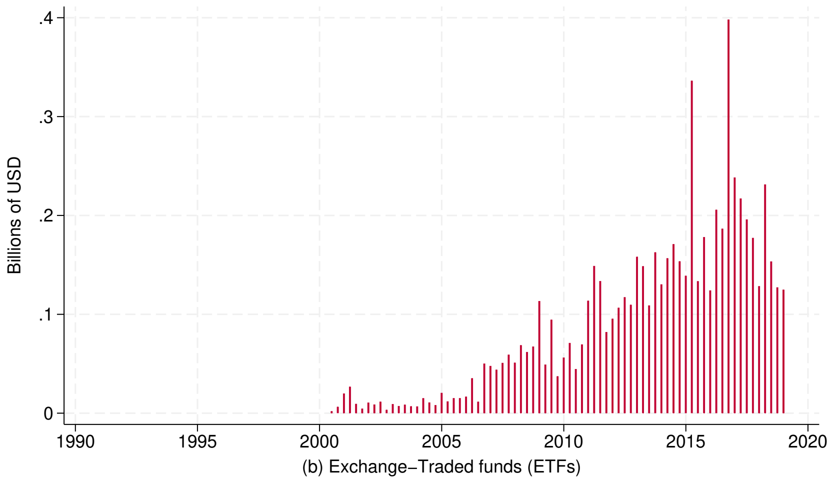
6 Tables and Figures

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 the ETF data, the sample is from 2000:Q1 to 2018:Q4



(a) Mutual Funds



(b) Exchange−Traded funds (ETFs)

Figure 2: **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 in the last quarter of five different years.

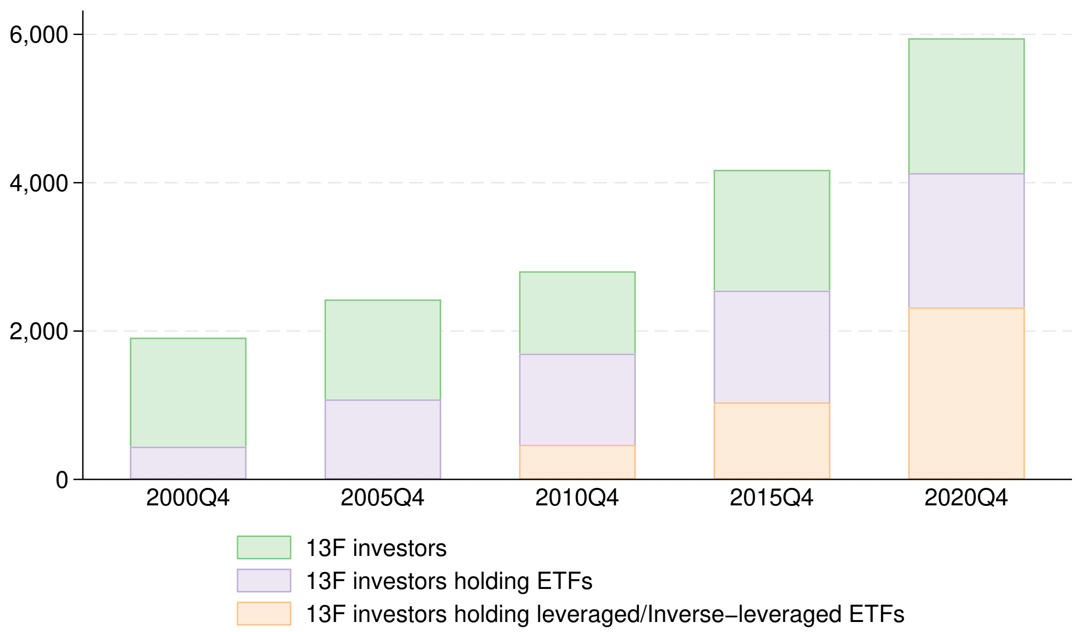
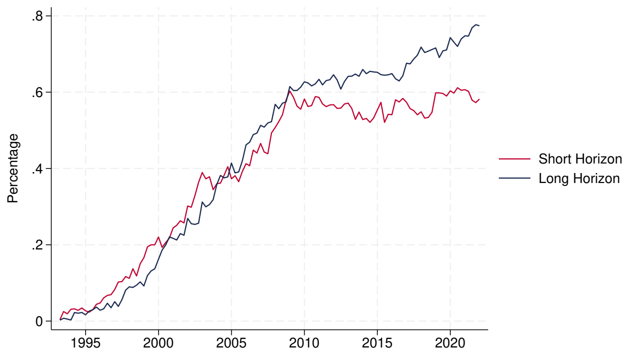
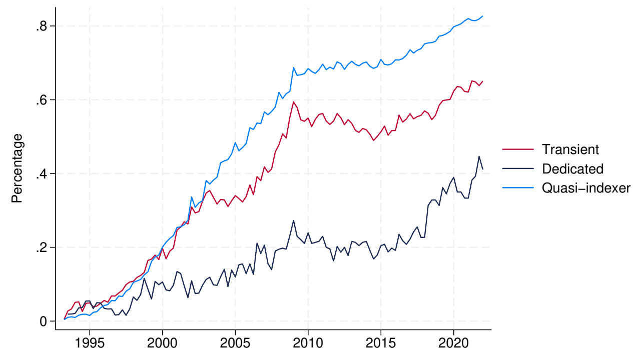


Figure 3: **The evolution in the adoption of ETFs in 13F Institutional Investors holdings**

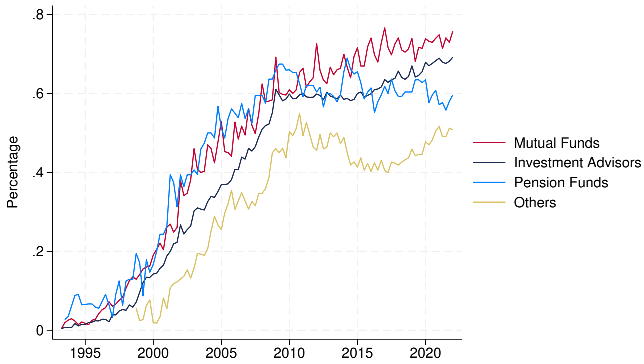
This figure shows the time series of the percentage of 13F institutional investors holding exchange-traded funds (ETFs) in their portfolios from 1993 to 2021. The 13F institutional investors are classified based on three criteria. In Panel A, investors are classified into short- and long-horizon based on the *average churn ratio* of Yan and Zhang (2009). In Panel B, we group investors into transient (i.e., those with high portfolio turnover and highly diversified portfolios), dedicated (i.e., those 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 according to 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)

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). The *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 the quarter end, in millions of US dollars. *Ownership* is the percentage of shares outstanding owned by all equity mutual funds (ETFs) in our sample. The *NYSE Decile* is the average NYSE size decile of a mutual fund(ETF) portfolio stock. Panel A reports the descriptive statistics for the sample of mutual funds. Panel B shows the results for the sample of exchange-traded funds (ETFs). Panel C reports the correlation coefficient for one-quarter (Q_{t-1}) to four-quarters (Q_{t-4}) lags in the number of owners. This is, the total number of mutual funds (ETFs) holding the same stock. Only stocks with market capitalization equal to or higher than 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	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

Panel C: Autocorrelation

	Mutual Funds				ETFs			
	Q_{t-1}	Q_{t-2}	Q_{t-3}	Q_{t-4}	Q_{t-1}	Q_{t-2}	Q_{t-3}	Q_{t-4}
Number of owners	0.861	0.851	0.786	0.78	0.832	0.807	0.791	0.701

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

This table reports the time-series statistics of cross-sectional 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 for 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 NYSE size decile 5 are included. The sample period for 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**

The standard deviation of daily stock returns over quarter t+1 (σ_{t+1}) is regressed on squared fragility $\sqrt{G}$ at quarter t and a set of lagged control variables as detailed in Equation (1) 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 by 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 (σ).

	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**

The standard deviation of daily stock returns over quarter *t+1* (σ_{t+1}) is regressed on squared fragility $\sqrt{G}$ at quarter t and a set of lagged control variables as detailed in Equation (1) 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 (σ).

	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)	(-.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: **Panel regression: Stock return volatility and fragility**

This table presents the results of a panel regression of the 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 define as *Controls FB* those control variables employed by Friberg et al. (2023) and include the log of market capitalization and the inverse of stock price. *Controls GT* refers to the control variables used by Greenwood and Thesmar (2011), which are 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 (σ). 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. The sample period is from 2009:Q1 to 2018:Q4.

	All firms				IO >0.2				Mkt cap >Median			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(11)	(11)	(12)
$\sqrt{G^{MF}}$	0.065***		0.032*	0.01	0.06***		0.046**	0.031	0.064		0.046***	0.034
	(3.60)		(1.86)	(0.70)	(3.37)		(2.57)	(1.56)	(3.58)		(2.59)	(1.61)
$\sqrt{G^{ETF}}$		0.187**	0.176**	0.152**		0.191**	0.179**	0.139**		0.193**	0.178**	0.147**
		(2.24)	(2.10)	(2.06)		(2.26)	(2.20)	(2.05)		(2.34)	(2.13)	(2.22)
Controls FB	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No
Controls GT	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Year-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	98,304	69,776	69,776	69,776	95,923	68,744	68,744	68,744	98,283	69,772	69,772	69,772
adj. R^2	0.662	0.683	0.689	0.725	0.661	0.683	0.688	0.711	0.662	0.683	0.691	0.748

Table 6: **Fragility and excess return volatility**

The standard deviation of *excess stock returns* over quarter *t+1* (σ^{exc}_{t+1}) is regressed on squared fragility $\sqrt{G}$ at quarter t. Excess returns are estimated based on the single-factor market model (1-Factor σ) the Fama and French (1993) three-factor model (3-Factor σ), and the Fama and French (1993) three-factor model augmented with the momentum factor of Carhart (1997) (4-Factor σ). 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. (1) based on ETF data only ($\sqrt{G}^{ETF}$).

Panel A: Mutual fund Fragility

	Full sample				2009 - 2018			
	1-Factor σ	3-Factor σ	4-Factor σ	DGTW	1-Factor σ	3-Factor σ	4-Factor σ	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 σ	3-Factor σ	4-Factor σ	DGTW	1-Factor σ	3-Factor σ	4-Factor σ	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 7: **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), (3), (4), and (5), we sum the ownership of the top 3 institutions, whereas for specifications (2), (6), (7), and (8), we take the top 10 institutions. The *bottom IO* represents the combined ownership of the smaller institutional investors whose equity holdings equal that of the top IO. The *middle IO* is the aggregated ownership of all institutional investors not considered in either the top or 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 the 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)	(3)	(4)	(5)	(6)	(7)	(8)
Top IO	0.471**	0.263**	0.568*	0.617**	0.530**	0.406***	0.424***	0.328**
	(2.71)	(2.37)	(5.00)	(4.37)	(3.50)	(4.29)	(4.44)	(3.40)
Mid IO	0.163**	0.184**	0.164**	0.115	0.100	0.158*	0.048	-0.064
	(2.23)	(2.06)	(2.06)	(1.32)	(0.89)	(1.75)	(0.46)	(-0.45)
Bottom IO	-0.466***	-0.157*	0.086	0.069	0.018	0.106	0.076	-0.039
	(-2.90)	(-1.75)	(0.72)	(0.58)	(0.13)	(1.08)	(0.72)	(-0.28)
G^{MF}	0.034***	0.022**	0.020**		0.019	0.025**		0.016
	(2.88)	(2.08)	(2.15)		(1.54)	(2.17)		(1.15)
G^{ETF}				0.308**	0.206**		0.288**	0.200*
				(2.25)	(1.98)		(2.17)	(1.90)
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	131,040	77,421	69,217	69,217	77,421	69,217	69,217
adj. R^2	0.659	0.667	0.652	0.689	0.689	0.652	0.689	0.703

Table 8: **Stock return volatility, ownership by large 13F institutional investors, and stock price fragility - alternative aggregation of institutional investors**

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), (3), (4), and (5), we sum the ownership of the top 5 institutions, while for specifications (2), (6), (7), and (8), we take the top 7 institutions. The *bottom IO* represents the combined ownership of the smaller institutional investors whose equity holdings equal that of the top IO. The *middle IO* is the aggregated ownership of all institutional investors not considered either in the top or 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 the 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 5 Inst	Top 7 Inst	Top 5 Inst			Top 7 Inst		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Top IO	0.467***	0.429***	0.652***	0.678***	0.616**	0.567***	0.594***	0.526**
	(3.35)	(3.45)	(5.62)	(5.84)	(4.01)	(5.65)	(5.83)	(4.38)
Mid IO	0.131*	0.125*	0.116*	0.053	0.038	0.095*	-0.003	0.024
	(1.87)	(1.74)	(1.70)	(0.52)	(0.79)	(1.79)	(-0.04)	(0.45)
Bottom IO	-0.284**	-0.227*	0.139	0.106	0.088	0.155	0.124	-0.029
	(-2.18)	(-1.91)	(1.42)	(1.00)	(0.83)	(1.60)	(1.20)	(-0.28)
G^{MF}	0.061***	0.052***	0.041*		0.031	0.037*		0.028
	(2.89)	(2.88)	(1.94)		(1.54)	(1.97)		(0.92)
G^{ETF}				0.284**	0.229**		0.244**	0.201*
				(2.33)	(2.01)		(2.16)	(1.99)
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	131,040	77,421	69,217	69,217	77,421	69,217	69,217
adj. R^2	0.659	0.667	0.652	0.689	0.689	0.652	0.689	0.703

Table 9: **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 covering 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 10: **Stock return volatility, excess return volatility, and activeness of ETFs**

This table presents the results of Fama and MacBeth (1973) regressions of next quarter's total return volatility and excess return volatility on the squared fragility of the current quarter. We estimate Fragility as detailed in Equation (1). 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 the 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 (σ). The sample period is from 2009:Q1 to 2018:Q4

	Total return volatility			Excesss return volatility					
	(1)	(2)	(3)	1-Factor σ	3-Factor σ	4-Factor σ	1-Factor σ	3-Factor σ	4-Factor σ
$\sqrt{G}^{ETF(Active)}$	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, ρ 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:

$$\frac{dP_t}{P_t} = \alpha dt + \sigma dZ_t \tag{1}$$

$$dB_t = rB_t dt, \tag{2}$$

where α 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 σ, 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 θ_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

$$\begin{aligned}
\mu_{it} &= \theta_i(\alpha - r) + r - \frac{c_{it}}{W_{it}} && (5)\\
q_i &= \begin{bmatrix}\theta_i\sigma & \sigma_{i,Liq}\end{bmatrix} && (6)\\
d\widetilde{Z}_i &= \begin{bmatrix}dZ_t & dZ_{i,Liq}\end{bmatrix}' && (7)
\end{aligned}$$

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, θ_{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 η_{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.

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

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

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

where $[\eta_1 \quad \eta_2]$ is a vector of asset ownership and Ω 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

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

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

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

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.*

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

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

OA2.Mutual Fund database construction procedure

1. Mutual Funds Holdings

   Following Pavlova and Sikorskaya (2023), to create the database of mutual fund holdings we use data from CRSP Mutual fund Database (CRSP, from June 2010 to December 2018) and Thomson Refinitiv S12 (TRS12, from March 1980 to December 2018). We mostly rely on CRSP data for the second part of the sample since its relatively more reliable and timely (Ben-David et al., 2023).

   - To merge both databases we employ the MFLINKS file.
   - As in Doshi et al. (2015) we first process TRS12 database and stay only with those observations where FDATE and RDATE are equal. To avoid employing stale data, we keep the first reported FDATE-FUNDNO observation per fund.
   - For the funds that report data more than once in a month, we keep the last reported information for that given month.
   - We use the MFLINKS file to include the WFICN identifier. Whenever the merging process produces non unique WFCIN-RDATE observations, we keep that with the highest assets.
   - We proceed to merge the S12type3 file with the holdings database to obtain CUSIP data that we use to include PERMNO from CRSP for each stock.
   - We split-adjust the shares variable.
   - to Include the CRSP holdings data we match WFICN to csrp fundno variable from MFLINKS.
   - Finally, we check for possible duplicated observations and keep the last reported data for each reported month.

2. Selecting Equity Mutual Funds

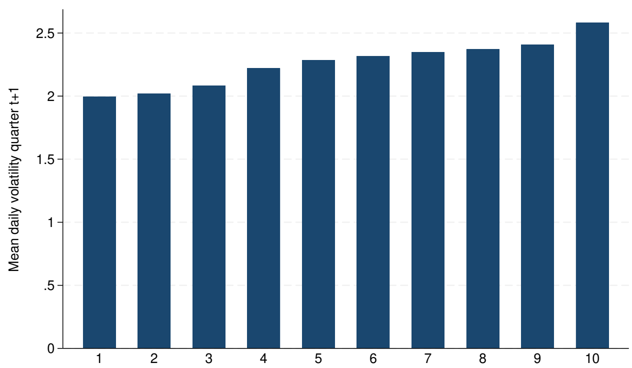
   - Based on the crsp obj cd variable, we exclude those funds whose names include: international, balanced, sector, bond, money market, and index.
   - Similarly, as with the holdings database, we keep the most recent entry for each fund.
   - If a fund changes its style during the sample period, we drop that fund from our sample.

OA3. Additional results

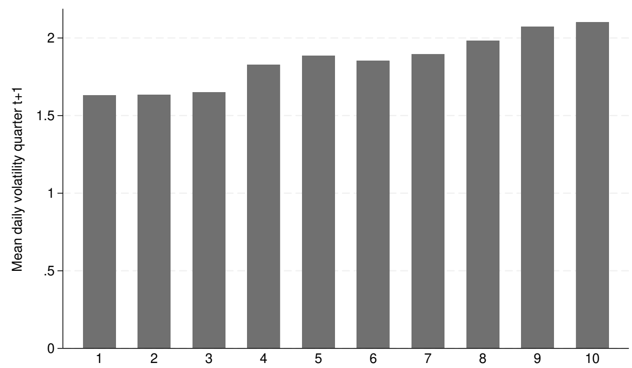
OA3.1 Volatility of fragility decile portfolios

Figure A1: **Fragility and volatility**

This figure shows, for each decile of Mutual fund and ETF fragility, the time series average of cross-sectional mean daily stock return standard deviation in the next quarter t+1. The sample covers the period from 1989:Q4 to 2018:Q4 for MF fragility deciles and from 2009:Q1 to 2018:Q4 for ETF fragility deciles.



MF fragility decile.



ETF fragility deciles

OA3.2 Stock characteristics of fragility quintile portfolios

Table A1: **Stock Characteristics**

For quarter t, stocks in our sample are sorted into 5 quintile portfolios based on their mutual fund (ETF) stock price fragility value. 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. *Volat* is the standard deviation of daily stock returns in the next quarter (t+1); *BM* is the book-to-market ratio; *Ret12* is the past 12-month stock return; *Turnover* is the average monthly share turnover (monthly volume traded over total shares outstanding) over the previous 3 months; *Age* is the firm's Age is calculated as the number of years (months/12) since the first return appears in CRSP; *Mkt 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. *NYSE* is the NYSE market capitalization decile breakpoint; *NOwn* is the average number of mutual funds (ETFs) that hold the same stock; *MOM* is the firm's stock return momentum decile; *Analysts* is the number of Analyst following the firm collected from I/B/E/S. Panel A shows the results for quintile portfolios sorted on fragility estimated as in Eq. (1) 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 for Panel A and from 2009:Q1 to 2018:Q4 for Panel B.

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

Quintile	Volat	BM	Ret12	Turnover	Age	Mkt cap	NYSE	NOwn	MOM	Analysts
1(low)	2.010	0.719	0.247	0.238	21.7	16,423.9	7.3	40.5	4.7	8.7
2	2.152	0.608	0.264	0.202	25.4	18,696.1	8.0	90.6	4.9	14.3
3	2.301	0.605	0.278	0.228	22.4	7,978.9	7.4	71.6	4.9	12.3
4	2.359	0.623	0.242	0.243	20.9	4,645.8	6.8	60.7	4.8	10.8
5 (high)	2.494	0.632	0.207	0.266	19.7	3,095.3	6.4	55.5	4.6	10.1

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

Quintile	Volat	BM	Ret12	Turnover	Age	Mkt cap	NYSE	NOwn	MOM	Analysts
1(low)	1.648	0.676	0.174	0.208	24.4	17,070.2	7.6	36.6	4.7	10.5
2	1.704	0.573	0.191	0.194	30.1	29,573.6	8.5	64.5	4.8	16.8
3	1.874	0.590	0.218	0.225	24.9	15,935.4	7.7	38.5	4.9	13.5
4	2.007	0.595	0.236	0.215	23.6	9,718.1	6.9	47.6	5.0	11.5
5 (high)	2.096	0.670	0.176	0.224	24.9	12,771.7	6.7	28.9	4.7	10.9

OA3.3 Fragility Return predictability

Daniel et al. (2001) provide a conceptual basis for employing mispricing measures as predictors of future returns. In their framework, the cross-section of expected security returns is determined by risk and misvaluation proxies.[1] As fragility essentially quantifies a firm's stock price exposure to non-fundamental demand shocks, it serves as a proxy for assessing the risk of misvaluation.[2] Based on the premise, We test ETF-based fragility ability to predict future stock returns in this section.

We start our empirical analysis by conducting univariate portfolio sorts using the estimated fragility measures derived from mutual funds and ETF data. Stocks within our sample are categorized into quintiles every quarter based on their mutual fund - G^{MF} (ETF - G^{ETF}) fragility from the preceding quarter. Subsequently, we construct long-short portfolios comprising the highest and lowest quintile-ranked stocks. Brown et al. (2021) report significant differences in ETF flows' return predictability across various time horizons, primarily noting that predictability persists for up to a 6-month horizon. In line with their findings, we create these long-short portfolios and track their excess return for one and two quarters into the future.

Panel A of Table A2 reports the return predictability of fragility-sorted equal-weighted portfolios. On average, for a holding period of one quarter, stocks that exhibit high Mutual fund (ETF) fragility (Quintile 1 - Q1) have a monthly excess re-

1 In Daniel et al. (2001) model, in equilibrium, securities are mispriced, and proxies for mispricing are informative about the future returns of different securities. Moreover, a fundamental assumption of the model is that some investors are overconfident about their abilities, which leads them to i) overestimate the quality of the information and their ability to interpret it and ii) may interpret noise as fundamental information. Arbitrageurs are able to exploit the pricing errors introduced by such noise traders by do not eliminate all mispricing due to risk aversion.

2 This interpretation of fragility has been employed in previous studies. For instance, Friberg et al. (2022) rely on stock price fragility as a proxy for exposure to mispricing and find that firms with higher fragility raise their cash holdings and lower their investments. The authors explore those results as evidence that firms respond to the risk of future misvaluation by leading corporate managers to take precautionary behavior by reducing their investments and increasing their cash holdings

turn of 0.895% (0.867%) while those stocks with the lowest fragility values (Quintile 5 - Q5) have a monthly excess return of 0.488% (0.544%). The spread portfolio (Q1-Q5) shows a monthly excess return of 0.407% (0.323%). If we extend the holding period to two quarters ahead, we observe that the spread portfolio shows a slightly higher excess return of 0.416% for mutual fund fragility portfolios and 0.382% for ETF fragility sorted ones. Overall, for equally-weighted portfolios, both fragility measures exhibit similar excess return. In panel C of Table A2, we address the concern that return predictability could stem from smaller firms in our sample by estimating value-weighted portfolio returns. We first notice that the average next quarter and next two-quarters excess return of the mutual fund fragility spread portfolio is considerably lower and no longer statistically different from zero. Second, we observe that for the ETF fragility sorted portfolios, the one-quarter excess monthly return of the spread portfolio is 0.245% (t-stat = 1.88) and 0.369% (t-stat = 2.17) for the portfolio that is held for two-quarters.

All return predictability from employing the mutual fund fragility measure loses statistical significance once we estimate value-weighted portfolio returns. This is not the case for ETF-fragility sorted portfolios. Moreover, we document a similar pattern as described by Brown et al. (2021) in which excess return is higher at a two-quarters horizon.

We also risk-adjust the equally-weighted and value-weighted spread portfolio (Q1-Q5) returns. We adjust excess returns 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 ofAmihud (2019) and the momentum factor of Carhart (1997) - FF5MA. Panel C of Table A2 shows that for equally-weighted returns, we observe positive alpha for both mutual fund and ETF-based fragility sorted spread portfolio for a one-quarter

Table A2: **Portfolio sorting on fragility: Excess returns and Abnormal returns**

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. Quintile 1 includes the stocks with the highest value of each stock price fragility measure, while Quintile Q5 is the lowest. Q1-Q5 is the spread portfolio that goes long Q1 and shorts Q5. Panel A presents the equal-weighted excess returns, while Panel B reports the value-weighted excess returns. Panel C reports the risk-adjusted returns for the equally-weighted spread portfolio (Q1-Q5) while Panel D shows the value-weighted risk-adjusted returns for that same portfolio. 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. Returns are in percent per month. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels. The sample period is from 2009:Q1 to 2018:Q4.

Panel A: Excess return Equally-weighted

	One quarter			Two quarters		
	Q1	Q5	Q1-Q5	Q1	Q5	Q1-Q5
G^{MF}	0.895*	0.488	0.407**	0.927**	0.511	0.416**
	(1.77)	(1.18)	(2.13)	(1.80)	(1.16)	(2.09)
G^{ETF}	0.867*	0.544	0.323**	0.934**	0.553	0.382**
	(1.93)	(1.37)	(1.98)	(2.09)	(1.39)	(2.18)

Panel B: Excess return Value-Weighted

	One quarter			Two quarters		
	Q1	Q5	Q1-Q5	Q1	Q5	Q1-Q5
G^{MF}	0.862*	0.658	0.204	0.750	0.562	0.188
	(1.83)	(1.58)	(1.32)	(1.61)	(1.54)	(0.96)
G^{ETF}	0.811*	0.566*	0.245*	0.932**	0.562	0.369**
	(1.91)	(1.67)	(1.88)	(2.27)	(1.55)	(2.10)

Panel C: Risk-adjusted return Equally-Weighted (Q1-Q5)

	One quarter			Two quarters		
	FF3	FF5	FF5MA	FF3	FF5	FF5MA
G^{MF}	0.289**	0.267**	0.263**	0.277**	0.188	0.185
	(2.40)	(2.14)	(2.01)	(2.22)	(1.47)	(1.41)
G^{ETF}	0.404***	0.372***	0.297**	0.502***	0.495***	0.376***
	(3.20)	(3.01)	(2.64)	(3.36)	(3.42)	(3.09)

Panel D: Risk-adjusted return Value-Weighted (Q1-Q5)

	One quarter			Two quarters		
	FF3	FF5	FF5MA	FF3	FF5	FF5MA
G^{MF}	0.106	0.103	0.014	0.023	0.246	0.008
	(0.78)	(0.75)	(0.09)	(0.23)	(0.56)	(0.07)
G^{ETF}	0.279**	0.277**	0.238**	0.393***	0.394***	0.279**
	(2.27)	(2.23)	(1.99)	(2.68)	(2.64)	(1.98)

horizon. However, the mutual fund-based portfolio shows non statistically significant alphas for FF5 and FF5MA asset pricing models for the two-quarter horizon. Panel C of Table A2 verifies previous findings and reports that, for the value-weighted portfolio, mutual fund fragility does not show statistically significant alphas for any risk-adjustment for both one and two quarters holding horizon. We observe the opposite for the case of ETF-based fragility sorted spread portfolio.

While our empirical analysis in this section is not exhaustive, it provides further evidence in line with the argument that relying on ETF can potentially improve our estimation of stock price fragility by examining its performance as misvaluation proxy. In a portfolio sorting approach, We find evidence of higher expected returns for a portfolio of high ETF fragility stocks compared to those with lower fragility. Moreover, the spread portfolio provides statistically significant excess returns and alphas. Our results are significantly weaker or non distinguishable different from zero in several cases if we rely on a mutual fund based fragility measure. Finally, our results align with those of Brown et al. (2021), documenting that non-fundamental demand shocks steaming from ETF flows have different effects at different time horizons.

OA4. Institutional investors leveraged/inverse ETF holdings

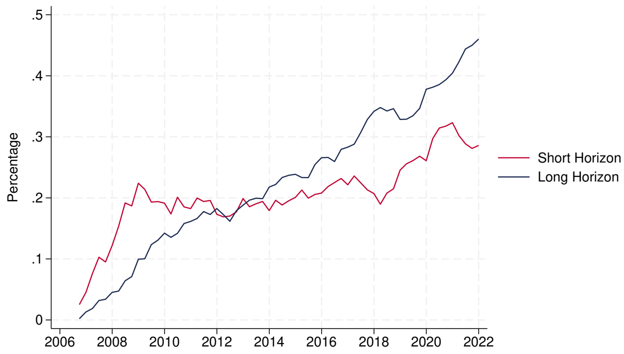
While the first ETF was launched on the Toronto Stock Exchange (TSX) in March 1990, the first US domestic ETF, the SPDR S&P 500, was introduced in January 1993. Since their inception, ETF attracted the attention of investors due to their hybrid design that combined characteristics of open and closed-end mutual funds while offering broad diversification at a lower cost and equity-like liquidity.

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 ETF 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 offer negative exposure by obtaining a negative multiplier of the benchmark return, for instance -1.5x -2x of the return.

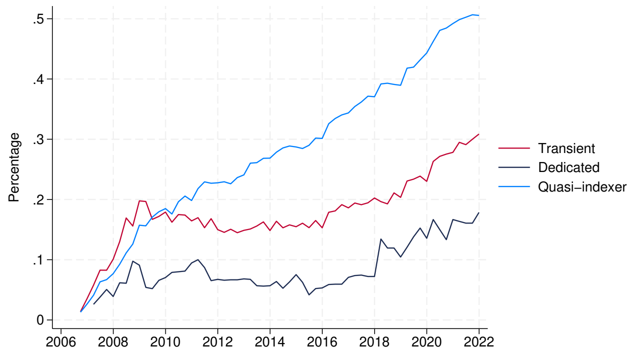
We identify leverage and inverse-leverage ETFs as those that include the following terms in their names: leverage, inverse, Double, Short, Ultra, UltraShort, 4x, 3x, 2.5x, 2x, 1.5x, 1.25x.

Figure A4: **13F Institutional Investors holding leveraged/inverse-leveraged ETFs**

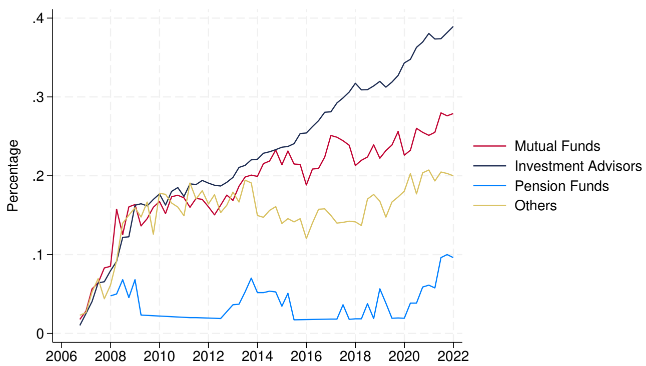
This figure shows the time series of the percentage of 13F institutional investors that held leverage or inverse-leveraged exchange-traded funds (ETFs) in their portfolios from 1993 to 2021. 13F institutional investors are classified based on three different criteria. In panel A, investors are classified 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)