ARTICLE IN PRESS

Journal of Financial Markets xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Financial Markets

journal homepage: www.elsevier.com/locate/finmar

An ETF-based measure of stock price fragility[☆]

Hamilton Galindo Gil[a,*], Renato Lazo-Paz[b]

[a] *Department of Finance and Economics at Monte Ahuja College of Business, Cleveland State University, 2121 Euclid Avenue, Cleveland, OH 44114, United States of America*

[b] *Telfer School of Management, University of Ottawa, 55 Laurier Ave E, Ottawa, Ontario K1N 6N5, Canada*

A R T I C L E I N F O

JEL classification:
G12
G14
G23

Keywords:
Non-fundamental demand risk
Fragility
Mutual funds
ETFs
Volatility

A B S T R A C T

Equity mutual fund flows are commonly employed to measure stock price fragility - a stock's exposure to non-fundamental demand risk. However, this approach may be biased by confounding fundamental information, potentially underestimating risk exposure. We propose an alternative method that uses the primary market data of exchange-traded funds (ETFs). This approach overcomes many limitations of mutual fund data, incorporates the influence of a broader set of investor demand, and strongly predicts stock return volatility and return comovement. Our study highlights the significant role that the arbitrage trading activity of ETFs play in signaling non-fundamental demand shocks.

1. Introduction

According to classical asset pricing theories, 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, previous papers document 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 (outflows) tend to underperform (outperform) in the long run (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 many researchers to use investor flows to and from mutual funds as sources of exogenous non-fundamental price pressure.[2]

☆ We would like to thank David Brown for sharing their list of exchange-traded funds. We are thankful for helpful feedback from Eser Arisoy, Markus Broman, Travis Box (discussant), Adelphe Ekponon, Eliezer Fich, Agustin Hurtado, Maksim Isakin, Fabio Moneta, Maurizio Montone, Pauline Shum Nolan, Qianru Qi, Francois-Eric Racicot, Matthew Ringgenberg, Yinjie Shen, Andrea Vedolin, Wei Wang, and seminar participants at the Bank of Canada, FMA 2023 Annual Conference, the AFA 2024 Ph.D. Poster Session, Cleveland State University, Nova School of Business and Economics (*PhD Pitch Perfect*), and the University of Ottawa.

* Corresponding author.
E-mail addresses: h.galindogil@csuohio.edu (H. Galindo Gil), rlaopaz@uottawa.ca (R. Lazo-Paz).

1 The effect of noise traders (De Long et al., 1990), trading based on informational and noninformational motives (Wang, 1996), and the limits to arbitrage activity (Shleifer and Vishny, 1997) on stock prices and trading volume, are modeled in some seminal theoretical papers.

2 Wardlaw (2020) and Berger (2022) discuss papers examining mutual fund flows as exogenous shocks to stock prices.

https://doi.org/10.1016/j.finmar.2024.100946
Received 29 January 2024; Received in revised form 2 October 2024; Accepted 3 October 2024

Available online 4 November 2024

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 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, empirical and theoretical studies have raised doubts about the validity of employing mutual fund flows as a proxy for non-fundamentally driven price pressure.

Recent empirical studies show that mutual fund flows motivate fund managers to perform discretionary trades[4] like choosing to sell low-quality stocks (Berger, 2022; Huang et al., 2022) and 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). In other words, even during extreme flows, such as fire sales, fund managers aim to minimize negative impacts by selectively choosing stocks and avoiding those more exposed. This contradicts the assumption that mutual fund managers simply expand or contract their portfolios in response to such flows, thus inducing price pressure in their trades. Additionally, Berk and Green (2004) in their theoretical work argue that mutual fund flows reflect investors' learning about mutual fund manager skills rather than solely indicating non-fundamental demand shocks. Overall, it is likely that the impact of mutual fund flows on prices cannot be exclusively attributed to non-fundamental demand and is, at most, a noisy proxy.

By contrast, relying on exchange-traded funds (ETF) flows provides three main advantages: (1) they signal observable non-fundamental demand shocks, (2) the process of ETF shares redemption and creation (i.e., flows) does not involve discretionary trades, and (3) the composition of ETF ownership captures demand from a broad cross-section of market participants.

The focus of this study is to provide an alternative method for estimating stock price fragility by employing data on ETFs. 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 increasing activeness of the ETF industry on asset prices (e.g., Easley et al., 2021; Davies, 2022).

Brown et al. (2021) argue that ETF flows are indicative of non-fundamental demand shocks. The authors show, theoretically, that the creation and redemption of ETF shares (i.e., ETF flows) mimics relative mispricing correction. Therefore, ETF premiums or discounts (i.e., relative mispricing) signal non-fundamentally driven price distortions. Empirically, if ETF flows signal the correction of non-fundamentally driven demand, it should be possible to predict future returns of both ETF shares and their underlying assets using ETF flows as indicators. This supports the notion that non-fundamental demand creates mispricing, which subsequently corrects. Consistent with their model and this hypothesis, Brown et al. (2021) find strong evidence of return predictability that persists over three- and six-month horizons. Furthermore, Davies (2022) finds similar empirical evidence in the context of leveraged and inverse-leveraged ETFs.[5]

Our analysis consists of two main parts. In the first part, we test the validity of our methodology by comparing its ability to forecast stock return volatility with that of the original estimation method. We start by estimating the stock price fragility measure as defined by 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 estimate the stock price fragility measure employing only ETF data, G^{ETF}. Finally, in a regression setting, we test the ability of each measure to forecast next-quarter daily stock return volatility. Additionally, we extend our empirical tests to the prediction of asset return comovement, by estimating the pair-wise level version of fragility, cofragility.[6]

In the second part of our analysis, we explore the factors that potentially make G^{ETF} a 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 return volatility (e.g., Gabaix, 2011; Ben-David et al., 2023) and whether increased ETF activeness (e.g., Easley et al., 2021; Davies, 2022) helps explain our results.

We highlight four main empirical results. First, we find that the statistical and economic significance of G^{MF} in forecasting next quarter's stock return volatility significantly declines in the second part of our sample (2009–2018) - out-of-sample. Greenwood and Thesmar (2011) document that from 1989 to 2008, an increase in G^{MF} fragility from the 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 G^{MF} 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 experienced significant outflows, as shown in Fig. 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 2018, approximately 70% of mutual funds and investment advisors in the 13F institutional investors holding database included ETFs in their portfolios.

[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., 2024), 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] Leverage ETFs were first introduced to the market in 2006. Similar to traditional ETFs, these funds offer exposure to a wide range of benchmarks; however, their replication method involves the use of derivatives. This mechanism allows ETF fund managers to leverage the fund's performance. While it is possible to

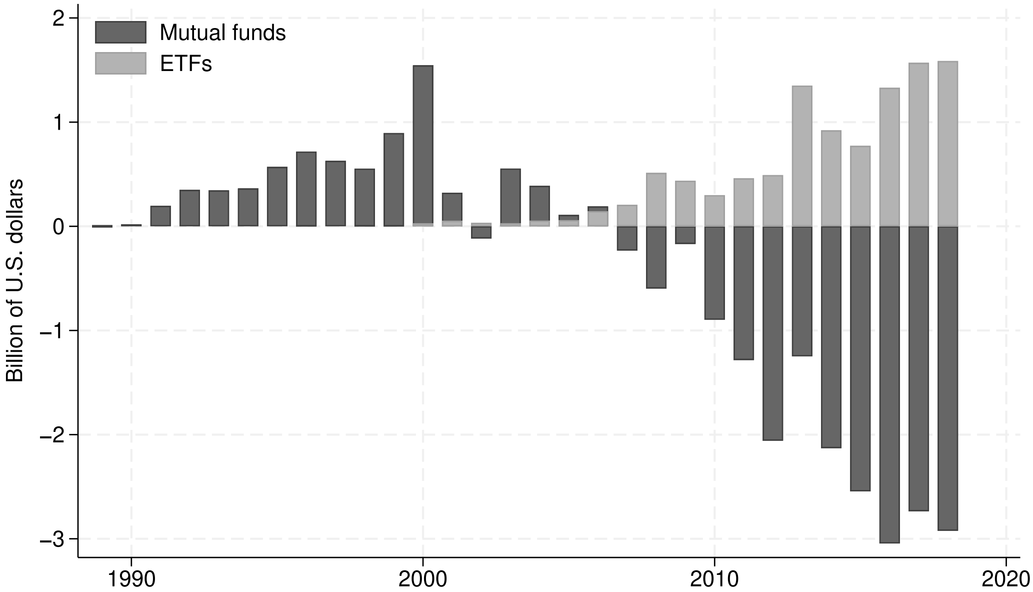

Fig. 1. Flows to equity mutual funds and exchange-traded funds (ETFs).
This figure illustrates the evolution of the total net cash flows, in billions of U.S. dollars, for our sample of equity mutual funds and exchange-traded funds (ETFs). The mutual fund data spans from the fourth quarter of 1989 to the fourth quarter of 2018, while the ETF sample covers the period from the first quarter of 2000 to the fourth quarter of 2018.

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). Specifically, its economic significance is six times larger compared to that of G^{MF}. A one-standard-deviation (SD) increase in G^{ETF} predicts an increase of 1.24% of a SD in return volatility, compared to a 0.22% increase predicted by a one SD increase of G^{MF} for the same sample period. Moreover, when both G^{MF} and G^{ETF} are included in our regression model, only the coefficient of G^{ETF} remains positive and statistically significant. This evidence supports the conjecture that G^{ETF} provides additional information on fragility beyond what is captured by G^{MF}. This evidence supports the conjecture that G^{ETF} provides information on fragility above and beyond that included in the G^{MF} measure. In an extension of this analysis, we find that an ETF-based co-fragility strongly predicts next-quarter stock return comovement, even when including several control variables and including current quarter covariance (correlation). Additionally, we document a significant decline in the predictive power of the mutual fund-based version on stock return comovement. Our evidence is consistent with 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} partially captures the influence of institutional investors' ownership on stock price volatility. In a recent study, Ben-David et al. (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 subunits 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 (Gabaix, 2011). In this context, 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. Additionally, we 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 for small- and mid-sized institutional ownership becomes statistically insignificant, and the coefficient for large institutional investors ownership is reduced. 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). Moreover, our results align with the literature on the impact of demand by institutional investors on asset prices (Bushee and Noe, 2000; Ben-David et al., 2021)

Fourth, we document that the forecasting power of G^{ETF} on the next quarter's stock price volatility is mostly explained by *active* ETFs. As is now widely documented, the growth of the ETF industry has been attributed to their provision of diversification at relatively lower costs, high intra-day liquidity, and superior tax efficiency compared to traditional mutual funds. However, the

achieve positive exposure (e.g., obtaining 1.5X or 2X the return of a specific benchmark), negative exposure is also feasible (inverse-leveraged). Investors can purchase ETFs that offer negative exposure by achieving a negative multiplier of the benchmark return, such as −1.5X or −2X.

[6] We thank the anonymous referee for suggesting this analysis.

evolution of the ETF industry has been marked by innovation that has blurred the distinction between the classic concepts of passive and active approaches to investment management (Easley et al., 2021). While large, low-cost, passive index ETFs are by far the biggest investment vehicles in terms of assets under management,[7] the growth of the ETF industry has been characterized by the development of specialized, industry-specific, and characteristic-based ETFs. Israeli et al. (2017), Davies (2022), and Ben-David et al. (2023) point out that the development of such investment products, broadly categorized as active ETFs, cater to investors' extrapolation beliefs and speculative demand, most likely reflecting sentiment driven demand. Thus, we hypothesize that more active ETFs are more likely to channel non-fundamentally driven demand shocks.

We begin exploring this hypothesis by estimating the activeness index of Easley et al. (2021) in 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. Next, we decompose the G^{ETF} into active and passive components following the methodology outlined by Greenwood and Thesmar (2011) and re-estimate our main findings. Our findings suggest that the primary driver of our results is the active ETFs component. These results are consistent with Ben-David et al. (2021), who suggest 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 our 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 support the hypothesis that ETF primary market flows provide valid signals of non-fundamental demand shocks (Brown et al., 2021), that institutional investors' ownership contributes to stock return volatility (Bushee and Noe, 2000; Koijen and Yogo, 2019; Ben-David et al., 2021), and that increased ETF activeness has likely attracted investors whose expectations about future returns are not necessarily warranted by fundamentals (Ben-David et al., 2023). 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 stock price fragility estimation method.

Our paper contributes to the ongoing discussion on using mutual fund flows as a proxy for non-fundamentally driven price variations. Consistent with prior research that point out the flaws in that empirical approach (Wardlaw, 2020; Huang et al., 2022; Berger, 2022), our findings indicate that the documented reduction in stock price fragility does not necessarily reflect a lower sensitivity to non-fundamental demand shocks. Instead, we suggest that the estimation of stock price fragility based on mutual fund flows might fail to adequately capture those shocks, whereas an ETF-based method proves more effective. This highlights potential limitations in the empirical reliability of mutual fund flow-based measures of fragility.

Additionally, our work adds to the growing literature on the effects of ETF activity on the volatility of their underlying assets. Ben-David et al. (2018) show that ETF trading induces non-diversifiable volatility due to short-horizon liquidity traders attracted by lower trading costs. Our results are complementary. We show that ETF ownership composition plays an important role in explaining those results. Koijen and Yogo (2019) find 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. Ben-David et al. (2021) shows that correlated trading activity among subunits of large and mid-sized institutional investors increases the difficulty of diversifying away shocks to their holdings, inducing higher volatility —a phenomenon known as granularity. We connect these findings to our results by showing that an ETF-based fragility measure partially captures the effect of institutional ownership on next-quarter stock return volatility. Thus, our approach effectively incorporates trading motives (i.e., flows), along with the effect of ownership composition, on non-fundamental volatility, thus providing a comprehensive approach to assessing firms' exposure to non-fundamental demand. Moreover, by providing a stock-level measure, we offer an empirical approach that facilitates further exploration of the heterogeneous impact of the trading and ownership of ETFs on their underlying assets.

Furthermore, our work joins earlier literature that argues that the development of the ETF industry has attracted noise trading. Israeli et al. (2017) argue that the features that make ETFs popular, such as intra-day liquidity and lower trading costs, also attract uninformed traders who might otherwise trade the underlying securities. Ben-David et al. (2023) demonstrate that ETF providers have responded to investor demand for specialized ETFs by launching thematic products that cater to trendy investment themes and investor sentiment. Similarly, Davies (2022) reports that leveraged ETFs effectively attract gambling-like demand. In line with this evidence, we show that the development of more active and specialized ETFs has likely drawn non-fundamental investor demand, as our results are primarily driven by the active ETFs in our sample.

Encouraged by the lack of theoretical models connecting economic foundations and fragility measures, we provide a theoretical model of stock price fragility in Appendix. Our goal is to link the microstructure of the economy, such as preferences, wealth, and non-fundamental shocks, to stock fragility. Specifically, based on Merton (1971), we propose a two-agent asset pricing model in which agents experience idiosyncratic liquidity shocks. In this framework, we show that changes in non-fundamental demand are determined by the volatility of the agents' liquidity shocks and, more importantly, the asset ownership structure.

Our model contributes to the asset pricing literature in two aspects. First, we attempt to formalize the definition of stock fragility in the model using structural characteristics of the economy, such as preferences and non-fundamental shocks, which are absent in the literature on stock fragility (Greenwood and Thesmar, 2011). Second, we consider risk-aversion heterogeneity in our model,

[7] For instance, according to ETF.com, by the end of 2018, the three largest ETFs – the SPDR S&P 500, the iShares Core S&P 500, and the Vanguard Total stock market – reported expense ratios between 0.3% and 0.9% and managed assets between 270 and 300 billion U.S. dollars.

which allows us to study the effect of wealth distribution on non-fundamental demand and, consequently, on stock fragility. This complements the heterogeneous-agent asset pricing literature by connecting investor heterogeneity and stock fragility, a topic not yet studied (e.g., Wang, 1996; Gârleanu and Panageas, 2015; Pohl et al., 2021). A heterogeneous-agent model that studies stock fragility based on mutual funds and ETFs seems promising, and we leave it for future work.

The remainder of this paper is organized as follows. In Section 2, we describe the conceptual framework supporting our empirical approach. In Section 3, we describe the mutual funds and ETF data sources. In Section 4, we present our main empirical results. We conclude in Section 5.

2. Conceptual framework

This section presents the theoretical framework underlying our empirical approach. We first review the literature on ETF primary market flows and non-fundamental demand shocks. Next, we examine studies linking firms' ownership structure to demand risk, connecting this to our methodology. Finally, we address the limitations of using mutual fund flows as a proxy for demand shocks, highlighting how ETF data can better estimate stock price fragility and resolve these issues.

2.1. *Non-fundamental demand shocks*

Non-fundamental demand shocks cause market participants to trade an asset with- out regard to fundamental information about changes in future growth prospects or risk factors, causing deviations in asset prices from their intrinsic values (De Long et al., 1990). These trades are categorized as noise/liquidity-driven (De Long et al., 1990; Wang, 1994) or sentiment-driven (Baker and Wurgler, 2006). Identifying such shocks is challenging due to the unobservable nature of fundamental values, and mutual fund flows have traditionally been used as a proxy, though recent studies (e.g., Berger, 2022; Huang et al., 2022) question this approach. We argue that ETF primary market flows provide clearer signals of non-fundamental demand shocks, which we explain through the ETF redemption/creation mechanism, building on Brown et al. (2021).

Consider two traded securities, A and B, which share the same fundamental value but have different prices ($p_A \neq p_B$). According to the law of one price, p_A should be equal to p_B. If this is not the case, *relative mispricing* arises. Moreover, it can be argued that $p_A - p_B$ captures the excess demand for either A or B. Thus, by taking the difference between A and B share prices, their shared fundamental value cancels out and the residual must reflect mispricing related to non-fundamental net excess demand. For efficiency to be restored, an arbitrageur should buy the less expensive asset and sell the more expensive one, generating price pressure on both. This arbitrage activity provides the signal required to identify and observe the non-fundamental demand shocks.

In the context of the ETF market, at $t = 0$ we assume that both the price of the ETF share (p_A) and the ETF basket of underlying assets (p_B) are the same ($p_A = p_B$). In the following period, $t = 1$, a shock of fundamental news and non-fundamental distortions leads to a difference in price between the two securities ($p_A \neq p_B$). At $t = 2$ Authorized Participants (APs) engage in arbitrage trading against mispricing to restore equilibrium prices. In the case of an ETF premium (i.e., ETF price above the underlying securities: $p_A > p_B$), APs will buy shares of the underlying basket of ETF securities and create and sell ETF shares. Conversely, if an ETF discount arises (i.e., ETF price below the underlying securities: $p_A < p_B$), APs will sell the underlying securities and redeem (i.e., buy) ETF shares. This redemption and creation activity generates price pressure for both ETF shares and the underlying assets, restoring equilibrium prices by $t = 3$. Thus, a very important highlight is that relative price efficiency is restored by affecting the supply of ETF shares.

Brown et al. (2021) argue that the temporary dislocation between the ETF's share value and the NAV of their underlying assets signals the appearance of a non-fundamental demand shock. Their empirical analysis supports this claim, showing that ETF flows can predict asset returns, which subsequently reverse, consistent with their theoretical predictions. Building on this, Davies (2022) introduces a Speculation Sentiment Index, revealing that speculative trades involving leveraged ETFs contribute to temporary distortions in market prices. Another appealing aspect of relying on ETF primary market flows is that unlike mutual fund flows, which might reflect investor learning about managerial skill (Berk and Green, 2004), ETF flows are driven by arbitrage opportunities exploited by APs, making them a more reliable indicator of market fragility since they more likely reflect price misalignments rather than skill-based decisions.

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) introduce this measure based on a model that represents changes in an investor's portfolio assets as a function of two key motivations: (1) those attributable to active rebalancing and (2) 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's fundamentals. If flow-driven demand cancels out across owners, prices should reflect only fundamental information. However, if non-fundamental demand is not resolved, 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) conclude that the two key determinants of a security's return variance due to non-fundamental demand are: (1) a vector representing the weight of each investor in that security (i.e., ownership concentration) and (2) the conditional variance–covariance matrix of flows originating from security owners (i.e., non-fundamental demand shocks).

More recently, Ben-David et al. (2021) expand the model of Greenwood and Thesmar (2011) to study the relationship between large institutional 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 et al. (2021) develop 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: (1) systematic aggregate shocks driving institutional trades; (2) fundamental idiosyncratic shocks; and (3) 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, theoretical models and empirical evidence suggest that non-fundamental stock return volatility is primarily driven by two key factors: ownership concentration and institutional ownership. Notably, institutional ownership is largely overlooked in the fragility measure of Greenwood and Thesmar (2011), as mutual funds are mainly held by households, while ETFs are owned by both institutional and retail investors (Dannhauser and Pontiff, 2019). We argue that an ETF-based fragility measure captures both effects due to this ownership split. The growing use of ETFs by institutional investors is well-documented,[8] as they use them for sector exposure, risk hedging (Huang et al., 2020), and circumventing 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 faces two key challenges: (1) identifying independent shocks to stock prices that are unrelated to firm fundamentals and fully observable, and (2) accessing comprehensive asset ownership data. The first challenge, explored extensively in the literature,[9] often uses mutual fund sales as a proxy for non-fundamental price shocks, as funds tend to sell off holdings during significant outflows (Coval and Stafford, 2007). Greenwood and Thesmar (2011) employ this approach to estimate stock fragility. However, recent studies have questioned key assumptions, such as proportional trading and the absence of discretionary trades.

For mutual fund flows to serve as an adequate proxy for exogenous price changes, non-fundamental demand shocks should be transmitted to all securities within the fund portfolio. This means mutual funds managers should adjust their positions proportionally in response to demand shocks, influencing the prices of all the underlying securities—an assumption known as proportional trading (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). Instead, they tend to avoid selling poorly performing, illiquid firms with lower growth. Thus, when empirically tested, the proportional trading assumption does not hold and might lead to biased inferences.

Closely tied to the proportional trading assumption is the assumption that mutual fund flows do not incorporate fundamental information from discretionary trades of fund managers. Huang et al. (2022) reveal that during fire sales, 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 suggests 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.

We argue that an ETF-based stock price fragility measure offers three key benefits for estimating fragility. First, it provides observable non-fundamental demand shocks. The arbitrage mechanism in the ETF primary market, through the creation and redemption process, offers a reliable and measurable signal of such shocks, which can be tracked by monitoring the number of outstanding ETF shares. Second, the approach eliminates reliance on assumptions about fund managers' behavior. The mechanical correction of misalignments between the ETF's NAV and its underlying assets reduces concerns about discretionary decisions that might introduce fundamental information into fund flows. Finally, this measure captures a broad range of market participants. The diverse ownership composition of ETFs allows it to reflect the influence of institutional demand on asset prices, a factor often overlooked by existing methodologies. Additionally, this broader ownership base might help address the second empirical challenge by providing more comprehensive data on investor ownership across a wider set of market participants.

[8] The Online Appendix details the increasing presence of ETFs in 13F institutional investors' portfolios, including data on the use of leveraged and inverse-leveraged ETFs. Our findings support previous research, demonstrating the extensive adoption of ETFs by institutional investors.

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

3. Data and variable construction

3.1. Mutual fund data

Our sample consists of U.S. mutual funds over the 1989–2018 time period. 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 in the last quarter of 1989 and continue to the last quarter of 2008. This allowed us to replicate their estimations (i.e., in-sample results). Second, although the first U.S.-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 2007–2009 (Madhavan, 2014). This period matches the end of Greenwood and Thesmar (2011) sample period. We examine the period starting in 2009 to test our measure, capturing the growth of ETF activity and applying an out-of-sample test of the original fragility measure in the context of increased passive investing.

We collect data on fund returns and total net assets (TNA) from the Center for Research in Security Prices (CRSP) Survivor-Biased-Free Mutual Fund database, We then obtain mutual funds' quarterly holdings data from the Thomson/Refinitiv Mutual Fund Database (*s12*). We merged the databases 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 U.S. domestic equity mutual funds and Pavlova and Sikorskaya (2023) to create the mutual funds holdings database.[10]

Consistent with previous studies, we restrict our holdings sample to include only stocks with market capitalizations at or above NYSE Size Decile 5. Greenwood and Thesmar (2011) highlight two advantages of applying this filter: (1) it simplifies matrix computations and (2) ensures that the estimation focuses on stocks of greater dollar importance more that are likely to be affected by liquidity-driven trades. Similarly, Francis et al. (2021) point out that an empirical issue in fragility estimation is that it becomes highly noisy if a stock has low mutual fund ownership, which is 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. Additionally, we exclude mutual funds with less than $ 5 million dollars in TNA.[11] Our fund sample includes 3871 distinct U.S. domestic equity mutual funds with 138,316 fund-quarter observations for the 1989–2018 period.

3.2. Exchange-traded funds data

Our ETF data sample covers the period from the first quarter of 2000 to the last quarter of 2018. To create our primary ETF database, we begin by reviewing the list of ETF identifiers from Brown et al. (2021).[12] To extend this database to include ETFs up to the last quarter of 2018, we follow their guidelines and combine ETFs listed in CRSP with the list pulled from http://www.etf.com/etfanalytics/etf-finder as of September 2019. In CRSP, we identify ETFs by selecting funds with a share code of 73

We combined our list of U.S. equity ETFs with information from Bloomberg and CRSP. From Bloomberg, we obtain data on outstanding shares and funds' 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 Survivor-Biased-Free Mutual Fund database data. We also exclude ETFs with less than $50 million in assets to limit the impact of illiquid and stale pricing due to infrequent trading.

We impose the same filters to stocks in our 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. In total, our sample includes 1096 distinct ETFs for which we have both holdings and price/return data.

In addition, we use stock-level data from additional sources: market data from CRSP, firm characteristics from Compustat, and Institutional Ownership data from Thomson/Refinitiv 13F institutional holdings database.

3.3. Estimating stock price fragility

We follow Greenwood and Thesmar (2011) and estimate stock price fragility, $G^{k_{it}}$, as follows:

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

where $\theta_{i,t}$ denotes the market capitalization of stock i at time t, $W_{i,t}^k$ denotes ownership concentration of stock i, and Ω_t^k is the conditional variance–covariance matrix of owners dollar flows in time t. K refers to either mutual funds or ETF data. Therefore, the two primary components of the fragility measure are a security's ownership composition and the variance–covariance matrix of its investors' non-fundamentally driven trades. An interesting aspect of this measure is that it provides a firm-level value that captures the combined effect of both components on non-fundamental risk. In other words, the higher the value of $G^{k_{it}}$, the greater a firm's exposure to non-fundamental flow-driven uninformed trading activities.

[10] We describe the merging of holdings databases and the selection of the mutual funds process in Section A.4 of the Appendix.

[11] While Greenwood and Thesmar (2011) do 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).

[12] We thank David Brown for providing us with these data.

The ownership structure is represented by a vector of each mutual fund (ETF) portfolio allocation to stock i relative to the fund's total net assets, as described in the following expression:

$$w_{i,k,t} = \frac{n_{i,k,t} P_{it}}{a_{k,t}}, \tag{2}$$

where $n_{i,j,t}$ denotes the number of shares of security i held by mutual fund (ETF) k at time t, P_{it} is the price of security i, and $a_{k,t}$ denotes the mutual fund (ETF) j total net assets (net asset value). Thus, vector $W^k_{i,t}$ summarizes the relative presence of all investors in stock i at time t.

3.3.1. Mutual fund-based fragility (G^{MF})

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

$$MFFlow_{k,t} = \frac{TNA_{k,t} - TNA_{k,t-1}(1 + R_{k,t})}{TNA_{k,t-1}}, \tag{3}$$

where $TNA_{j,t}$ denotes the mutual fund k TNA for quarter t and $R_{k,t}$ denotes 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^{MF}_t to be expressed in dollar terms. We follow Greenwood and Thesmar (2011) and rescale the Ω_t matrix by fund assets at time t to obtain an estimate $\hat{\Omega}_t$:

$$\hat{\Omega}^{MF}_t = diag(TNA_t)\Omega^{MF}_t diag(TNA_t), \tag{4}$$

For each quarter t, we calculate $\hat{\Omega}^{MF}_t$ using a five-year rolling window estimation starting from the first quarter of 1984. We then estimate G^{MF} as detailed in Eq. (1).

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 portfolio allocation weights of ETFs to each stock i multiplied by the stock's i price and divided by the NAV of ETF k. In the context of the ETF primary market, we calculate *ETF flows* as the percentage changes in the number of shares outstanding for each ETF k at each time t:

$$ETFFlow_{k,t} = \frac{SharesOutstanding_{k,t}}{SharesOutstanding_{k,t-1}} - 1. \tag{5}$$

We next normalize the ETF fund flows covariance matrix $\Omega_{k,t}$ as follows:

$$\hat{\Omega}^{ETF}_t = diag(NAV_t)\Omega^{ETF}_t diag(NAV_{k,t}). \tag{6}$$

To ensure consistency with the mutual fund-based fragility estimation process, we estimated $\hat{\Omega}^{ETF}_t$ using a five-year rolling window.[13] Before 2005, ETF holdings represented only a negligible percentage of a stock's outstanding shares (Da et al., 2020). Consequently, using 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 report ETF-based fragility values from 2009 onwards. This approach guarantees the inclusion of a more substantial dataset and helps mitigate the potential for noisy results. Finally, we calculate G^{ETF} as described in Eq. (1).

Fig. 1 shows the evolution of the total new cash flows, in billions of U.S. dollars, to our sample of mutual funds and ETFs from 1989 to 2018. In the early sample period, pre-2006, mutual funds mostly experienced inflows. However, beginning in 2006, mutual funds on aggregate experienced outflows. By contrast, ETFs experienced significant inflows over the years, especially in the later part of the sample period. Our results are consistent with those of Easley et al. (2021), who documents that active equity mutual funds in aggregate experienced net outflows while ETFs attracted net inflows. Interestingly, the authors find that the largest inflows of funds have gone to active ETFs, indicating an increasing investor preference for those investment vehicles.

Table 1 presents summary statistics of our mutual funds (Panel A) and ETF (Panel B) samples. In any given year, our sample includes more mutual funds (1134) than ETFs (334). The average ETF is larger in terms of AUM and holds a larger number of stocks. This difference is most likely driven by the presence of very large broad-based ETFs.[14] 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 of U.S. equity over time (Da and Shive, 2018; Glosten et al., 2021). Specifically, ETF ownership increased from an average of 0.63% from 2000 to 2009 to 3.96% from 2010 to 2018. Regarding the relative size of the underlying securities, the median stock for mutual funds is in the 8th NYSE decile, while for ETFs, it is in the 7th decile.

[13] 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.

[14] 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 AUM of $216.4, $253.4, and $337.2 billion, respectively.

Table 1
Summary statistics.
This table reports summary statistics for our sample of mutual funds and exchange-traded funds (ETFs). *Number of funds* is the average 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 summarizes the descriptive statistics for our sample of mutual funds. Panel B shows the results for the ETFs sample. 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 in our sample. For the mutual fund sample, the period spans from the last quarter of 1989 to the last quarter of 2018. For the ETF sample, the period extends from the first quarter of 2000 to the last quarter of 2018.

Panel A: Mutual Funds								
	Full Sample					Mean by period		
	Mean	Std	p25	Median	p75	1989–1999	2000–2009	2010–2018
Number of funds	1138	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
NAV (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

According to Greenwood and Thesmar (2011), for fragility to reliably forecast future volatility, a firm's ownership composition should remain relatively stable from one quarter to the next. We test this assumption by estimating the autocorrelation coefficient of the number of owners, defined as the number of mutual funds (ETFs) holding 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, 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). Additionally, we provide evidence that the ownership structure is also highly persistent for our ETF sample. This requirement on the persistence of ownership of sample firms means that if we observe a fund's ownership of stock i in quarter t, we require that same stock i to be part of the fund's 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. Moreover, this principle applies to more active ETFs, such as smart-beta ETFs. These funds typically deviate from their benchmarks by altering their weighting schemes rather than changing the selection of stocks they hold (Madan, 2010; Easley et al., 2021). These results provide additional support for the suitability of ETF data for estimating stock price fragility.

Table 2 provides a summary of the cross-sectional distributions of the variables that compose the fragility measure, as well as for the square root of G^{MF} and G^{ETF}.[15] Panel A shows that the number of mutual funds and ETFs holding the same stocks increased over time, particularly in the ETF sample for the latter part of our sample period. On average, stocks within the mutual fund (ETF) sample are held by 50 (25) funds.

The results in Panel B of Table 2 provide 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 exhibit an increase in the initial segment 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.[16] By contrast, the volatility of ETF flows experienced a substantial increase over the entire sample period, particularly in 2014–2018.

We also observe some interesting patterns concerning the correlation between fund flows, as presented in Panel C of Table 2. In particular, we document a decrease in the mean flow correlation for the mutual funds and ETFs samples. However, after an initial drop, the correlation among ETF flows remains fairly stable from 2009 to 2018. We also find that the flow correlation in both the

15 Greenwood and Thesmar (2011) prefer to use the square root of fragility because it is proportional to variance. In this way, fragility can also be defined as the conditional expected variance of non-fundamentally driven trades.

16 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; thus, effectively reducing fund's exposure to flow volatility.

Table 2
Fragility and fragility components descriptive statistics.
This table presents 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 mutual funds (ETFs) holding the same stock. *Flow volatility* represents the standard deviation of mutual (ETF) fund flows. *Flow correlation* is the Pearson correlation of mutual fund (ETFs) flows at the fund-pair level for each quarter. *Fragility* (sqrt) refers to the square root of the mutual-fund-based and ETF-based fragility measures, estimated as in Eq. (1). Only stocks whose market capitalization is equal to or higher than NYSE size decile 5 are included. For the mutual fund sample, the full period spans from the last quarter of 1989 to the last quarter of 2018. For the ETF sample, that period extends from the first quarter of 2000 to the last quarter of 2018. The mutual fund (ETF) 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

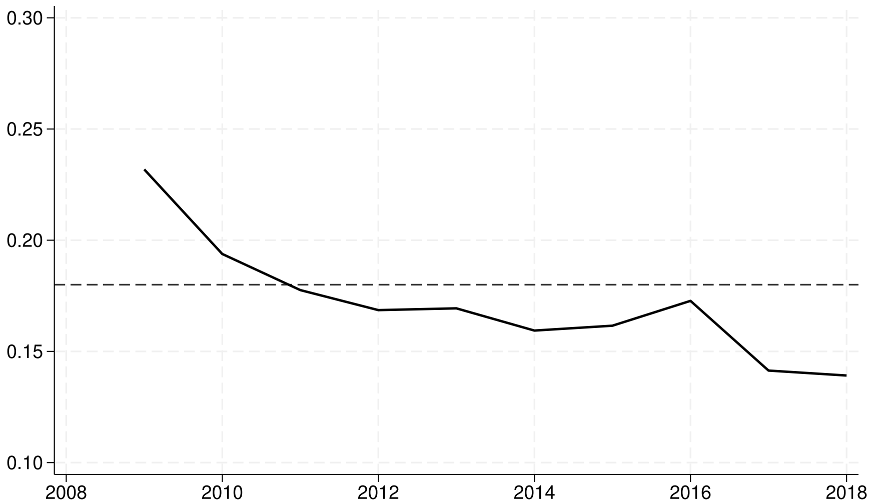

Fig. 2. Correlation between G^{MF} and G^{ETF}.
This figure shows the time-series yearly correlations between G^{MF} and G^{ETF}. The dotted red line indicates the average correlation over the sample period from the first quarter of 2009 to the last quarter of 2018. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

bottom (p25) and top (p75) quintiles is quite similar for mutual fund flows and ETF flows throughout the entire sample period. Regarding the $\sqrt{G^{MF}}$ variable, as shown in Panel D of Table 2, its mean value increased significantly from 0.039 to 0.143 during the period from 1989 to 2009, consistent with the pattern documented by Greenwood and Thesmar (2011); However, from 2009 to 2018 this value declined, averaging 0.102. This represents an approximate 30% decrease in the mean value during the out-of-sample period. By contrast, the mean $\sqrt{G^{ETF}}$ value continued to rise steadily for that same period.

Fig. 2 shows the time series yearly correlation between the G^{MF} and G^{ETF} variables. We estimate the correlation between these two measures using firm-level observations from each year between 2009 and 2018. The dotted line represents the time-series average of the cross-sectional correlations over the sample period, which is 0.18. Notably, the correlation between these measures shows a decreasing trend suggesting that although both measures initially capture a common component, its significance diminishes over time. Therefore, it is interesting to test whether both measures can coexist in our data. Our objective is to determine whether G^{ETF} can provide additional information beyond that provided by G^{MF}, thereby better capturing the effects of non-fundamental demand shocks on stock return volatility. We will explore this possibility in the next section.

A potential concern is that fragility values may be affected by differences in the characteristics of stocks included in the mutual fund and ETF samples. To address this, Table A1 in the Online Appendix presents an analysis where stocks are sorted into quintile portfolios based on G^{MF} (Panel A) and G^{ETF} (Panel B). We then calculate the time-series averages of cross-sectional means for various stock characteristics. Our findings, consistent with Greenwood and Thesmar (2011), show that fragility values do not consistently correlate with the number of owners, stock turnover, past returns, and analyst coverage. Interestingly, while we confirm that smaller firms and growth stocks with lower B/M ratios exhibit higher G^{MF} values, we do not find the same pattern when examining the quintiles in the G^{ETF} sample. Additionally, in figure A2 of the Online Appendix, we confirm (Greenwood and Thesmar, 2011) findings and observe a clear and positive correlation between G^{MF}, G^{ETF} and subsequent stock price volatility. Overall, we do not find a significant monotonic relationship between stock characteristics and fragility values, reducing concerns about a common underlying factor driving the correlation between fragility and stock return volatility.

4. Empirical results

In this section, we present our primary analysis to validate 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. We also extend this analysis to predict return comovement. Additionally, we expand our initial setting to incorporate the influence of institutional investors' ownership on stock price volatility and explore its implications for the ETF-based fragility measure. Finally, we consider the heterogeneity of the ETF industry by decomposing ETF-based fragility to examine the roles of active and passive ETFs in our 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) cross-sectional regression[17]:

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

where $\sigma_{i,t+1}$ is the one-quarter-ahead standard deviation of daily stock returns. $Z_{i,t}$ denotes 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, the lagged skewness of stock returns, the log of firm's age (in months), and share turnover. In several regression specifications, we also include the current standard deviation of daily stock returns ($\sigma_{i,t}$), Mutual fund (ETF) ownership, the log of the number of owners, and the Herfindahl index as controls. Fragility is winsorized at the 1% and 99% levels. 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 reports the results of the Fama-MacBeth regression. We begin by regressing future volatility on $\sqrt{G}^{MF}$ across the entire sample period. The results are presented in the columns (1)-(4). 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 ETF-based fragility measure, we run the same regression specifications on $\sqrt{G}^{ETF}$ for the same period.

Consistent with previous findings, in Table 3, we find a positive and statistically significant relationship between $\sqrt{G}^{MF}$ and next-quarter daily return volatility (Greenwood and Thesmar, 2011; Francis et al., 2021; Friberg et al., 2024). Nonetheless, it is worth noting that the reported coefficient is considerably smaller than the value documented by Greenwood and Thesmar (2011), who report a value for β of 0.696. In our analysis, we find this coefficient to be 0.459. Furthermore, when focusing on the latter part of our sample period, 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, as previously documented by Sias (1996) and Bushee and Noe (2000), and that the explanatory power of $\sqrt{G}^{MF}$ extends beyond pure ownership concentration, as proxied by the Herfindahl index. In column (4), we assess whether the predictive power of $\sqrt{G}^{MF}$ remains robust when accounting for a comprehensive set of control variables, including lagged volatility, since, as it is widely documented that volatility is highly persistent over time. The results reveal that the coefficient on $\sqrt{G}^{MF}$ decreases significantly to 0.072 (t-stat = 2.75). Moreover, if we focus on the latter part

[17] We perform (Fama and MacBeth, 1973) regressions to control for the effect of common trends like the increasing ownership of mutual funds and ETFs (Da et al., 2020).

Table 3
MF and ETF fragility and stock return volatility.
This table reports the results of cross-sectional Fama-MacBeth regressions of the standard deviation of daily stock returns over quarter *t+1* (σ_{t+1}) on squared fragility $\sqrt{G}$ at quarter *t* and a set of lagged control variables as detailed in Eq. (7). Fragility is measured by employing only mutual fund flows and holdings data ($\sqrt{G}^{MF}$) and ETF data only ($\sqrt{G}^{ETF}$). *IO* is the mutual fund (ETF) ownership (i.e., number of shares held divided by the number of shares outstanding). *log(numb owners)* is the log of the number of mutual fund (ETF) owners. *Own Herfindahl* is the ownership Herfindahl index H. The additional 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 firm's age, share turnover, and the lagged dependent variable (σ_t). ***, **, * denotes statistical significance at the 10%, 5%, and 1%. Standard errors are based on Newey–West adjustments. *t*-statistics are reported in parentheses. G^{MF} and G^{MF} are winsorized at the 1% and 99% levels. For the mutual fund sample, the *full sample* spans from the last quarter of 1989 to the last quarter of 2018.

	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
Obs.	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

of our sample, as shown in column (8), 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 previous analysis employing the $\sqrt{G}^{ETF}$ measure 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 of 0.825 (t-stat = 7.76), which is significantly higher than that of 0.325 (t-stat = 8.75) for $\sqrt{G^{MF}}$ for the same period. The results in column (10) 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 of 0.338 (t-stat = 5.93)

Overall, this initial analysis provides evidence that an ETF-based fragility measure is useful in forecasting next quarter volatility. Moreover, our estimates indicate that the original measure of Greenwood and Thesmar (2011) has lost forecasting power over time. This suggests that relying solely on the mutual fund-based measure might significantly underestimate a firm's exposure to non-fundamental demand shocks.

Next, we investigate an ETF-based fragility is a robust measure of non-fundamental demand risk and whether it captures information beyond what the original measure provides. Table 4 presents the results of an analysis of the volatility predictors for the later part of the sample period. We report the results of the Fama-MacBeth 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 compared to mutual funds, being held nearly equally by households and institutional investors. Therefore, it can be anticipated that the effect of non-fundamentally driven demand captured by $\sqrt{G}^{ETF}$ differs from that of $\sqrt{G}^{MF}$.

To explore this hypothesis, we repeat the analysis reported in Table 3 including both $\sqrt{G}^{ETF}$ and $\sqrt{G}^{MF}$ simultaneously. While we observe that daily volatility is positively and statistically significantly correlated with both measures, the coefficient of $\sqrt{G}^{ETF}$, 0.790 (t-stat = 7.77), is significantly larger than that of $\sqrt{G}^{MF}$, 0.067 (t-stat = 1.99). 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. To address the robustness of our findings, we include the full suite of controls and the lagged dependent variable. We report these results in column (4). The coefficient on $\sqrt{G}^{MF}$, 0.009, is no longer statistically different from zero (t-stat = 1.03). By 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 and suggests that $\sqrt{G}^{ETF}$ better captures a firm's exposure to nonfundamental risk.

To ensure robustness, we re-estimate the regression specification for column (4) for a set of subsamples based on firm's size and industry affiliation, and we find results that remain statistically significant. First, since for comparability we excluded firms with market capitalizations below the fifth NYSE size breakpoint (Greenwood and Thesmar, 2011), we group firms into mid-cap and large-cap categories. Mid-sized firms are those included in the NYSE 5th, 6th, and 7th deciles, while large-cap firms include those in

Table 4
MF and ETF fragility and stock return volatility: Subsamples by stock size and industry.
This table reports the results of cross-sectional Fama-MacBeth regressions of the standard deviation of daily stock returns over quarter *t+1* (σ_{t+1}) on squared fragility $\sqrt{G}$ at quarter t and a set of lagged control variables as detailed in Eq. (7). Fragility is measured by employing only mutual fund flows and holdings data ($\sqrt{G}^{MF}$) and ETF data only ($\sqrt{G}^{ETF}$). IO^{MF} (IO^{ETF}) is the mutual fund (ETF) ownership. The additional 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 firm's age, share turnover, and the lagged dependent variable (σ). Results are reported for the full sample (columns (1) through (4)), and for subsamples grouped by firm size (columns (5) and (6)) and industry sectors (columns (7) through (10)). We define mid-cap stocks as those included in the 5th, 6th, and 7th deciles of the NYSE size breakpoints. Similarly, we categorize large-cap stocks as those included in the 8th, 9th, and 10th deciles. We group stocks into industry sectors based on the Fama–French 12 industry classification. To ensure sufficient observations for each regression, we require that a sector includes a minimum of the 10% of the total observations. ***, **, * denotes statistical significance at the 10%, 5%, and 1%. Standard errors are based on Newey–West adjustments. t-statistics are reported in parentheses. G^{MF} and G^{MF} are winsorized at the 1% and 99% levels. The sample period is from the first quarter of 2009 to the last quarter of 2018.

					Size		Industry			
					Mid-cap	Large-cap	Manuf	Buss Eq	Wholesale	Financ
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\sqrt{G}^{MF}$	0.067*		0.015	0.009	0.079	0.016	0.045	0.037	0.120	0.084
	(1.99)		(1.16)	(1.03)	(1.15)	(0.60)	(0.57)	(0.46)	(1.32)	(1.18)
$\sqrt{G}^{ETF}$	0.790***		0.795***	0.426***	0.312***	0.587***	0.325**	0.348*	0.429**	0.510***
	(7.77)		(8.20)	(7.95)	(5.67)	(6.17)	(2.21)	(1.86)	(2.02)	(2.60)
IO^{MF}		0.014***	0.012***	0.005***	0.004***	0.004***	0.008***	0.002*	0.005**	0.007***
		(11.11)	(12.37)	(7.47)	(5.67)	(5.19)	(2.77)	(1.84)	(2.82)	(6.34)
IO^{ETF}		0.002**	0.012***	0.007***	0.007***	0.007***	0.008**	0.007*	0.009***	0.008**
		(2.03)	(6.58)	(4.96)	(5.17)	(4.81)	(3.42)	(1.92)	(5.77)	(2.25)
Add Controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	44,956	44,956	44,956	44,956	25,261	19,695	4601	4420	4818	6102
Adj. R^2	0.015	0.025	0.034	0.376	0.421	0.496	0.480	0.438	0.497	0.423

the 8th, 9th, and 10th deciles. Interestingly, we observe that the coefficient on $\sqrt{G}^{ETF}$ is larger for large-cap firms (larger stocks). Our results align with the liquidity trading hypothesis of Ben-David et al. (2018), who show that arbitrageours tend to concentrate their portfolios on the largest stocks in the ETF portfolios to minimize their transaction costs. This, in turn, helps explain why larger stocks are more exposed to shocks from the ETF market.

Second, following Huang et al. (2020), we group stocks into industry sectors based on the Fama–French 12 industry classifications. To ensure our regressions have enough observations, we require that each sector includes at least 10% of the total observations in our sample.[18] Our results show an economically and statistically significant relationship between $\sqrt{G}^{ETF}$ and next-quarter stock volatility for firms in the Manufacturing, Business Equipment, Wholesale, and Financial services sectors. Overall, our results confirm that G^{ETF} strongly predicts stock return volatility, and this effect is not limited to specific firm sizes or industries.

To address concerns that the regression settings might influence our results, we present panel fixed effects estimates, aligning with previous studies (Ben-David et al., 2021; Friberg et al., 2024). We include firm and year-quarter fixed effects and adjust the standard errors for clustering at the firm level. The sample period for this analysis spans from the first quarter of 2009 to the last quarter of 2018. We follow Friberg et al. (2024) specification and test the relationship between stock price fragility and future return volatility within three subsets: (1) the full sample; (2) a subset comprising observations with a minimum of 20% institutional ownership; and (3) 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.

The results in columns (1), (5), and (9) of Table 5 closely match those of Friberg et al. (2024).[19] We observe that $\sqrt{G}^{ETF}$ is positive and statistically significant in all three subsets, as shown 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. (2024) (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 our ETF-based fragility measure is a robust and strong predictor of future return volatility.

[18] Huang et al. (2020) find that most industry ETFs are concentrated among 10 of the 12 Fama–French industries, with only 121 ETFs classified as industry ETFs during the same sample period we employ. Given this reduced number of observations and to ensure our analysis covers a significant portion of industries while including enough observations for reliable estimations, we impose a minimum required number of observations.

[19] The disparities we observe might potentially be attributed to differences in sample periods since (Friberg et al., 2024) conduct their analysis from 2001 to 2017. In untabulated results, we replicate our analysis for their same time frame and obtained results that closely mirror those reported.

Table 5

MF and ETF fragility and stock return volatility: Panel regressions.

This table presents the results of a panel regression of the standard deviation of daily stock returns over quarter *t+1* (σ_{t+1}) on squared fragility $\sqrt{G}$ at quarter t and a set of lagged control variables. Fragility is measured by employing only mutual fund flows and holdings data ($\sqrt{G}^{MF}$) and ETF data only ($\sqrt{G}^{ETF}$). *Controls FB* include those control variables employed by Friberg et al. (2024), such as 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 firm's age, share turnover, and the lagged dependent variable (σ). Results are reported for the full sample (columns (1) through (4)), and for subsamples that include only firms whose mutual fund (ETF) institutional ownership (IO) is above 20% (columns (5) through (8)) and firms whose market capitalization is above the sample median (columns (9) through (12)). t-statistics are reported in parentheses and are based on standard errors clustered at the stock level. ***, **, * denotes statistical significance at the 10%, 5%, and 1%. G^{MF} and G^{MF} are winsorized at the 1% and 99% levels. The sample period is from the first quarter of 2009 to the last quarter of 2018.

	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

Lou (2012) is among the first to propose a capital-flow-based explanation for some return predictability patterns. Specifically, 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 themselves.[20] 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-factor 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).

In Panel A of Table 6, our results corroborate those of Greenwood and Thesmar (2011) as we observe that the coefficient of $\sqrt{G}^{MF}$ is slightly smaller than that obtained when analyzing total return volatility. Furthermore, there is 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, shown in the second part of Panel B. 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.

To summarize, the analysis in this section provides evidence supporting the argument that an ETF-based fragility measure strongly predicts future stock return volatility. Such predictive power holds for several subsamples and goes beyond several determinants of future volatility. Additionally, our findings align with previous studies that suggest ETF primary market flows signal non-fundamental demand shocks. While the ETF-based measure may capture a similar component to the MF-based fragility, namely, retail investors' demand, it may also incorporate the influence of a broader set of investors' demand.

4.2. Cofragility and comovement

Next, we expand our empirical analysis to the prediction of asset return comovement. To this purpose, we implement the co-fragility measure developed by Greenwood and Thesmar (2011). According to the authors, similar to the fragility measure, the covariance in returns between a pair of stocks is influenced by two components: one that reflects news about the fundamentals of both securities and another driven by correlated, non-fundamentally driven trades and common ownership. Since the fragility measure captures the latter component, a natural extension is to estimate a pair-wise measure to predict this covariance. The co-fragility measure is estimated as follows:

$$G_{ijt} = \frac{W_{it}\Omega_t W_{jt}}{\theta_{it}\theta_{jt}}, \tag{8}$$

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

Table 6
MF and ETF fragility and excess return volatility.
This table reports the results of cross-sectional Fama-MacBeth regressions of the standard deviation of *excess stock returns* over quarter *t+1* (σ_{t+1}^{exc}) on squared fragility $\sqrt{G}$ at quarter *t*. Fragility is measured by employing only mutual fund flows and holdings data ($\sqrt{G}^{MF}$) and ETF data only ($\sqrt{G}^{ETF}$). 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). Additionally, we also include as dependent variable returns adjusted Following Daniel et al. (1997) (DGTW), which adjusts returns variations originating from the size, book-to-market, and momentum stock characteristics. ***, **, * denotes statistical significance at the 10%, 5%, and 1%. Standard errors are based on Newey–West adjustments. *t*-statistics are reported in parentheses. G^{MF} and G^{MF} are winsorized at the 1% and 99% levels. For the mutual fund sample, the *full sample* spans from the last quarter of 1989 to the last quarter of 2018.

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

where G_{ijt} denotes the co-fragility between stocks *i* and *j* at time *t*. To predict correlations, we need to adjust the co-fragility measure by the square roots of each stock's fragility, as follows:

$$G_{ijt}^{s} = \frac{G_{ijt}}{\sqrt{G_{it}G_{jt}}}. \tag{9}$$

We next estimate the co-fragility measure using mutual fund data (G_{ijt}^{MF} and $G_{ijt}^{s,MF}$) and ETF data ($G_{ijt}^{ETF} and G_{ijt}^{s,ETF}$). For comparability, we follow Greenwood and Thesmar (2011) procedure and limit our sample to the largest 500 stocks with positive mutual fund (ETF) ownership in each quarter.

To test the co-fragility ability to forecast covariance and correlation, we perform the following Fama-MacBeth forecasting regressions:

$$\sigma_{ijt+1} = \alpha + \beta\sqrt{G_{ijt}} + \delta Z_{ijt} + u_{ijt+1}.$$

Similarly, for the case of pairwise correlations between stocks *i* and *j*, we run the following regression:

$$\rho_{ijt+1} = \alpha + \beta\frac{G_{ijt}}{\sqrt{G_{it}G_{jt}}} + \delta Z_{ijt} + u_{ijt+1},$$

where σ_{ijt+1} (ρ_{ijt+1}) denotes the covariance (correlation) between the daily returns of stocks *i* and *j* over the quarter *t+1*. We account for the determinants of return covariance (correlation), Z_{ijt}, by controlling for stock-pair-level fundamental characteristics. Specifically, the difference in NYSE and book-to-market (BE/ME) size deciles between each stock-pair. Additionally, we incorporate industry dummy variables to identify firms within the same two, three, and four-digit Standard Industrial Classification (SIC) codes. To account for the effect of common ownership on stock return comovement (Antón and Polk, 2014; Da et al., 2020), we include the logarithm of one plus the number of common mutual fund or ETF owners. Consistent with previous analyses, we run our regressions for the sample period from the first quarter of 2009 to the last quarter of 2018. We report those results in Table 7. Our results provide strong evidence of the ETF-based co-fragility measure's (G_{ijt}^{ETF}) ability to predict stock return comovement, as compared to the mutual fund-based measure (G_{ijt}^{MF}).

We next test the univariate relationship between the (G_{ijt}^{MF}) and (G_{ijt}^{ETF}) co-fragility measures and next-quarter stock return covariance. The results in columns (1) and (4) indicate that both measures have predictive power. However, when control variables are included, the relationship between covariance and G_{ijt}^{MF} is significantly attenuated and turns statistically indistinguishable from zero (*t*-stat = 1.11) when the current quarter covariance is added to the regression model, as shown in column (3).

In sharp contrast, the predictive power of G_{ijt}^{ETF} remains economically and statistically sizeable after including control variables and the current quarter covariance, as shown in column (5). In this specification, a one-standard-deviation increase in ETF co-fragility

Table 7

MF and ETF cofragility, covariance and correlation.

This table reports the results of Fama-Macbeth regressions of next-quarter (*t+1*) covariance (correlation) in daily stock returns between stocks *i* and *j* on Co-Fragility - $\sqrt{G_{ijt}}$ - (Co-Fragility rescaled - $G_{ijt}/\sqrt{G_{it}G_{jt}}$) and a set of control variables. In columns (1) through (6), the dependent variable is the one-quarter-ahead covariance between daily returns of pairs of stocks *i* and *j*. In columns (7) through (12), the dependent variable is the next-quarter correlation between daily returns of stocks *i* and *j*. Co-Fragility and Co-Fragility rescaled measures are calculated employing only mutual fund flows and holdings data (G^{MF}), and ETF data only (G^{ETF}). The control variables include the difference in NYSE size deciles (*size*), the difference in book-to-market (BE/ME) deciles between stocks *i* and *j*. We also account for common ownership as the log of one plus the number of common owners. To account for industry-wide effects, we include dummy variables as equal to one when stock *i* and stock *j* belong to the same two-digit ($SIC2_{it} = SIC2_{jt}$), three-digit($SIC3_{it} = SIC3_{jt}$) or four-digit($SIC4_{it} = SIC4_{jt}$) SIC code. In columns (3), (6), (9), and (12), we additionally control for current quarter covariance (σ_t) and correlation (ρ_t). ***, **, * denotes statistical significance at the 10%, 5%, and 1%. Standard errors are based on Newey-West adjustments. *t*-statistics are reported in parentheses. G^{MF} and G^{ETF} are winsorized at the 1% and 99% levels. The sample period is from the first quarter of 2009 to the last quarter of 2018.

	Dependent variable = σ_{ijt+1}						Dependent variable = ρ_{ijt+1}					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\sqrt{G^{MF}_{ijt}}$	0.637**	0.326*	0.199									
	(1.99)	(1.79)	(1.11)									
$\sqrt{G^{ETF}_{ijt}}$				2.094***	1.665**	1.088**						
				(2.74)	(2.34)	(2.01)						
$G^{MF}_{ijt}/\sqrt{G^{MF}_{it}G^{MF}_{jt}}$							0.008***	0.006**	0.004*			
							(2.98)	(2.21)	(1.70)			
$G^{ETF}_{ijt}/\sqrt{G^{ETF}_{it}G^{ETF}_{jt}}$										0.028***	0.021***	0.019***
										(7.18)	(5.74)	(6.47)
SIC2 = SIC2		0.113***	0.068***		0.138***	0.086***		0.069***	0.043***		0.071***	0.047***
		(6.21)	(5.07)		(8.15)	(8.72)		(17.62)	(17.05)		(13.95)	(16.57)
SIC3 = SIC3		0.251***	0.147***		0.223***	0.126***		0.051***	0.037***		0.034***	0.031***
		(6.86)	(6.14)		(5.35)	(4.78)		(6.11)	(4.43)		(3.60)	(2.82)
SIC4 = SIC4		0.177***	0.097**		0.168**	0.104**		0.096***	0.052***		0.105***	0.062***
		(4.04)	(2.37)		(3.07)	(2.24)		(13.35)	(5.99)		(14.00)	(7.04)
σ_t			0.437***			0.421***						
			(9.77)			(8.38)						
ρ_t									0.390***			0.380***
									(22.03)			(20.20)
Add Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Obs.	2,769,055	2,670,943	2,479,938	2,107,581	2,079,938	1,982,002	2,769,055	2,670,943	2,466,452	2,107,581	2,079,938	1,982,002
Adj. R^2	0.010	0.028	0.214	0.001	0.029	0.200	0.001	0.033	0.181	0.005	0.044	0.186

leads to a 0.026% increase in return covariance, which is about 35% of the sample covariance standard deviation. The fact that G^{ETF}_{ijt} holds its forecasting power in all regression settings supports the hypothesis that an ETF-based fragility measure is informative in predicting stock return comovement.

We find similar results for predictive regressions on future stock return correlations. The results in column (8) and (9) show that the normalized value of $G^{s,MF}_{ijt}$ keeps its predictive power on the next quarter stock return correlation even after controlling for additional variables and current quarter correlation, respectively. However, the regression coefficient (0.004) is marginally statistically significant at 10% (t-stat = 1.70). By contrast, $G^{s,ETF}_{ijt}$ shows a much stronger predictive relationship (0.019) that remains statistically and economically significant (t-stat = 6.47) after accounting for our full set of confounding factors, as shown in column (12). Here, a one-standard deviation increase in $G^{s,ETF}_{ijt}$ forecasts a 4.5% increase in return correlation, whereas a one-standard deviation increase in $G^{s,MF}_{ijt}$ predicts only a 1.1% increase.[21]

Taken together, the findings provide strong support for the hypothesis that an ETF co-fragility measure has predictive power on next-quarter stock return correlation and covariance. Moreover, this predictive power is significantly greater than that derived from the mutual fund co-fragility measure. Our results are consistent with the evidence suggesting that ETF arbitrage activity and secondary market trading increase return comovement (Broman, 2016; Da and Shive, 2018) and align with recent findings indicating that the rise of the ETF industry has led to an increase in the correlation among stock returns, possibly limiting diversification benefits (Fang et al., 2023).

4.3. Fragility and institutional investors' ownership

In this subsection, we explore potential determinants of the superior forecasting power G^{ETF} on stock return volatility and return comovement. We argue that a potential explanation is that ETF primary market flows channel excess demand from a broader cross-section of market participants, including both retail and institutional investors. In contrast, mutual fund flows primarily reflect the activity of retail investors.

---

[21] We estimate an average 15.4% pair-wise correlation in our sample period from the first quarter of 2009 to the last quarter of 2018. Our results are significant larger than those reported by Campbell et al. (2001) who reported an average correlation of 5.7% for the period between 1962 to 1997.

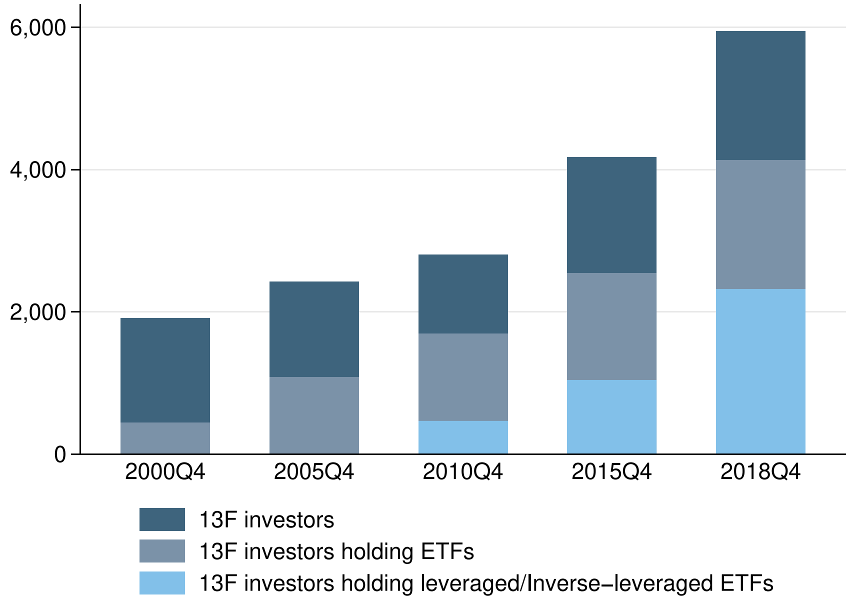

Fig. 3. 13F Institutional Investors including ETFs in their holdings.
This figure shows the total number of 13F institutional investors, as well as the subset of those investors holding ETFs and leveraged/inverse-leveraged ETFs in their portfolios. The data are presented for the last quarter of five different years.

Fig. 3 provides a snapshot of the total number of institutional investors included in the 13F database and the proportion of those investors who included ETFs in their holdings at five different points during our sample period. The figure shows a significant increase in the number of 13F institutional investors over the years, alongside 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. By the end of 2018, this proportion had surged to about 70%. Interestingly, a similar exponential growth is observable in the use of leverage and inverse-leverage ETFs.

In Fig. 4, we further analyze the time-series adoption of ETFs in 13F institutional investors' holdings. Additionally, we group 13F data by different investor types, revealing that investment advisors, mutual funds, quasi-indexers, and transient institutions are among those that have most extensively incorporated ETFs into their portfolios over our sample period.[22] 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 by long-term investors. Our findings align with the widespread inclusion of ETFs in 13F institutional investors' holdings.

In a recent study, Ben-David et al. (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 (Gabaix, 2011). 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.

Previous studies show 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.

To test the hypothesis that an ETF-based measure of stock price fragility effectively captures the impact of institutional investor trading, we follow Ben-David et al. (2023) framework. We conduct panel regressions, incorporating variables that account for ownership by large-, mid-, and small-sized institutional investors based on their AUM. Our objective is to assess how these variables influence the predictive power of G^{MF} and G^{ETF} on future return volatility. Our main specifications estimate is as follows:

$$\sigma_{i,t+1} = \beta_1 \mathrm{TopIO}_{i,t} + \beta_2 \mathrm{MidIO}_{i,t} + \beta_3 \mathrm{BottomIO}_{i,t} + \delta Z_{i,t} + \beta_4 G_{i,t} + \alpha_i + \theta_t + \mu_{i,t}, \tag{10}$$

[22] We use three common institutional investor classifications. First, Bushee (2001) categorizes institutions as transient (high turnover), dedicated (low turnover with concentrated holdings), and quasi-indexer (low turnover, well-diversified). Second, Yan and Zhang (2009) groups investors into short-horizon and long-horizon based on churn ratio. Finally, Koijen and Yogo (2019) classifies investors into mutual funds, investment advisors, pension funds, and others.

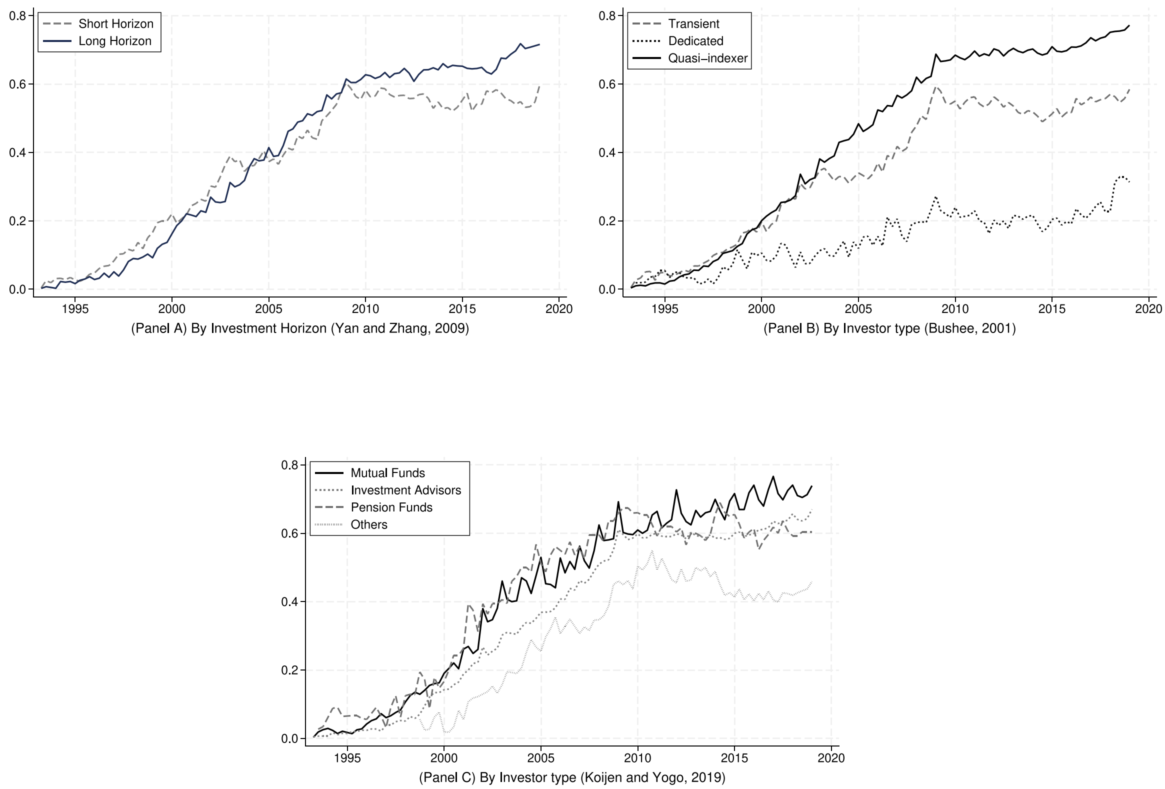

Fig. 4. The evolution in the adoption of ETFs in 13F Institutional Investors holdings.
This figure shows the time series of the percentage of 13F institutional investors holding exchange-traded funds (ETFs) in their portfolios from 1993 to 2018. The 13F institutional investors are classified based on three criteria. In Panel A, investors are classified into short- and long-horizon based on the average churn ratio as defined by Yan and Zhang (2009). In Panel B, we group investors into transient (i.e., characterized by high portfolio turnover and broad diversification), dedicated (i.e., exhibit large investments in portfolio firms, as well as low portfolio turnover), and quasi-indexer (i.e., characterized low portfolio turnover but maintains more diversified portfolios) (Bushee, 2001). In Panel C, we group investors according to Koijen and Yogo (2019) classification. The 13F holdings data are obtained from Thomson/Refinitiv, while ETF data is collected from Bloomberg and CRSP.

where $\sigma_{i,t+1}$ denotes the next quarter t stock i volatility. $\text{TopIO}_{i,t}$ denotes the fraction of shares outstanding collectively held by the top institutions ranked based on the money value of portfolio holdings over the previous four quarters. $\text{BottomIO}_{i,t}$ represents the aggregate stock's i ownership of the smallest institutional investors whose aggregate money holdings value equals that of the top institutions. $\text{MidIO}_{i,t}$ denotes collective ownership by institutions not classified as top neither as bottom. $Z_{i,t}$ denotes 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 denotes the stock fixed effect, and θ_t is the time (calendar year-quarter) fixed effect. The error term is denoted by $\mu_{i,t}$, and the standard errors are adjusted for clustering at the firm and time levels. Table 8 presents the results of these regressions, grouping the largest institutional investors as the top 3 and top 10 largest institutions.

Columns (1) and (2) report the estimation results for specifications that only include institutional investors' ownership by size. Our results closely follow those reported by Ben-David et al. (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 only large investors affect volatility.[23]

Ben-David et al. (2021) argue that including stock price fragility has little impact on the analysis of institutional investors granularity because each variable 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. We test this hypothesis by including G^{MF} in the regression specifications. Consistent with the findings of Ben-David et al. (2021), we observe in columns (3) and (6) that the coefficients for large- and mid-sized institutional ownership remain positive, while those for small institutional ownership are negative, both statistically significant. Additionally, G^{MF} positively predicts next quarter volatility. We next rerun this analysis but replace G^{MF} with G^{ETF}. We find

[23] Under this logic, when ownership is atomistic (i.e., each institution owns smaller portfolios), institutions can easily find multiple trading counterparties. Consequently, any price impact due to demand shocks is limited and effectively diversified away, thereby reducing its impact on stock volatility.

Table 8

MF and ETF fragility and stock return volatility: By ownership of 13F institutional investors.

This table reports panel regressions of next quarter's (*t+1*) daily stock return volatility (σ_{t+1}) on Institutional Ownership aggregations, fragility measures from mutual fund data (G^{MF}) and ETF data (G^{ETF}), and control variables. Volatility is calculated as the standard deviation of daily stock returns within each quarter. Following Ben-David et al. (2021), institutional investors are classified by size (top, mid, bottom) based on a rolling four-quarter average of their equity holdings. *Top IO* represents the largest institutional investors' aggregate ownership; specifications (1), (3), (4), and (5) sum the top 3, and (2), (6), (7), and (8) sum the top 10. *Bottom IO* represents smaller investors with equivalent equity holdings, while *Middle IO* includes the rest. Control variables include the Amihud (2002) illiquidity measure, stock price inverse, book-to-market ratio, market capitalization, and 6-month momentum. *t*-statistics (in parentheses) use clustered standard errors. ∗∗∗ , ∗∗, and ∗ indicate 1%, 5%, and 10% significance levels. G^{MF} and G^{ETF} are winsorized at 1% and 99%. The sample covers Q1 2009 to Q4 2018.

	2009–2018		2009–2018					
	Top 3 Inst	Top 10 Inst	Top 3 Inst			Top 10 Inst		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Top IO	0.752**	0.528**	0.568*	0.617**	0.503**	0.406***	0.424***	0.328**
	(3.71)	(3.92)	(5.00)	(4.37)	(3.50)	(4.29)	(4.44)	(3.40)
Mid IO	0.241***	0.168***	0.164**	0.115	0.100	0.158*	0.048	−0.064
	(2.34)	(3.76)	(2.06)	(1.32)	(0.89)	(1.75)	(0.46)	(−0.45)
Bottom IO	−0.114**	−0.314***	−0.086*	0.069	0.018	−0.106*	0.076	−0.039
	(−2.10)	(−4.10)	(−1.72)	(0.58)	(0.13)	(−1.78)	(0.72)	(−0.28)
G^{MF}			0.020**		0.019	0.025**		0.016
			(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.	77,421	77,421	77,421	69,217	69,217	77,421	69,217	69,217
Adj. R^2	0.585	0.595	0.652	0.689	0.689	0.652	0.689	0.703

Table 9

MF and ETF fragility and stock return volatility: Alternative aggregation of 13F institutional investors.

This table presents the results of panel regressions of next quarter's (*t+1*) daily stock return volatility (σ_{t+1}) on a set of different aggregations of Institutional Ownership, stock price fragility estimated based on mutual fund data only (G^{MF}) or ETF data only (G^{ETF}) and a set of control variables. We estimate stock volatility as the standard deviation of daily stock returns within each quarter. Following Ben-David et al. (2021), we classify institutional investors by size (top, mid, and bottom) based on a rolling four-quarter average of the rankings of their aggregate equity holdings.*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 (in parentheses) use clustered standard errors. ∗∗∗ , ∗∗, and ∗ indicate 1%, 5%, and 10% significance levels. G^{MF} and G^{ETF} are winsorized at 1% and 99%. The sample covers Q1 2009 to Q4 2018.

	2009–2018		2009–2018					
	Top 5 Inst	Top 7 Inst	Top 5 Inst			Top 7 Inst		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Top IO	0.734***	0.619***	0.652***	0.628***	0.506**	0.567***	0.514***	0.426**
	(4.33)	(3.09)	(5.62)	(5.84)	(4.01)	(5.65)	(5.83)	(4.38)
Mid IO	0.177**	0.125*	0.116*	0.053	0.038	0.095*	−0.003	0.024
	(2.77)	(1.99)	(1.70)	(0.52)	(0.79)	(1.89)	(−0.04)	(0.45)
Bottom IO	−0.284**	−0.227*	−0.139*	0.026	0.088	−0.155*	0.024	−0.029
	(−2.18)	(−1.96)	(−1.82)	(0.99)	(0.83)	(−1.79)	(1.20)	(−0.28)
G^{MF}			0.041*		0.031	0.037*		0.028
			(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.	77,421	77,421	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

in columns (4) and (7) that the coefficient on large institutional ownership and G^{ETF} are positive and statistically significant, respectively. However, the coefficients of the mid and small institutional ownership are smaller and indistinguishable from zero.

We next include both G^{MF} and G^{ETF} in our main regression model and find results in columns (5) and (8) similar to those documented previously. That is, the G^{MF} coefficient (0.019 and 0.016) loses statistical significance (t-stat = 1.54 and 1.15), while

Table 10
Activeness of ETF sample.
This table reports the time-series averages of the cross-sectional mean, median, standard deviation (*Std*), and 90th percentile (*P90*) of the activeness index (%). We estimate the activeness index (%) following the methodology defined by Easley et al. (2021) as the aggregated difference of the weight of each stock in each ETF and the corresponding weight of that stock in the value-weighted portfolio of all US-listed stocks (i.e., the market portfolio). We report the descriptive statistics for the sample period covering the first quarter of 2009 to the last quarter of 2018 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

the G^{ETF} coefficient (0.206 and 0.200) remains economically and statistically significant (t-stat = 1.98 and 1.90). For robustness, we replicate the previous analysis for alternative groupings of top institutional investors, specifically into top 5 and top 7 institutions, as shown in Table 9. Our results remain qualitatively the same.

Our evidence supports the hypothesis that the 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.

4.4. ETF activeness and stock price volatility

In this subsection, we analyze another likely advantage of our ETF-based measure: its ability to capture the impact of the increasing activeness of the ETF industry on asset prices. In principle, purely "passive" investing is typically associated with buy-and-hold strategies involving market index funds. While broad-based ETFs fit this traditional definition of passive investment, the ETF industry has evolved by launching products that blur the line between active and passive investing (Easley et al., 2021). Although these developments have driven the rise of the ETF industry, previous research suggests that the features that initially made ETFs popular, along with the emergence of more active and specialized ETFs, have attracted the attention of uninformed and sentiment-driven investors.

We begin by analyzing the activeness level of our ETF sample. Easley et al. (2021) propose a measure to identify active ETFs. According to the authors, active ETFs are those constructed to generate alpha (active-in-form) or used by investors as components of active portfolios (active-in-function). Since we aim to capture the influence of both types of active ETFs, we follow their methodology and estimate the Activeness Index as follows:

$$ActivenessIndex_{i,t} = \sum_{s=1}^{N} w_{i,s,t} - w_{market,s,t}, \tag{11}$$

where $w_{i,s,t}$ denotes the weight of stock i in fund s portfolio at time t, and $w_{market,s,t}$ denotes the weight of stock s in the value-weighted portfolio of all U.S. listed stocks (i.e., the market portfolio) at time t. By comparing ETF holdings to the market portfolio, this measure captures both activeness in function and form. It does so by contrasting ETF holdings with a completely passive investment strategy, which involves investing in all U.S.-domiciled stocks with weights proportional to their market capitalizations.[24] For an all-equity fund that has no leverage or short position, the activeness index lies between 0 an 1 and indicates the fraction of the fund's holdings that differ from the broad-based market portfolio.

We categorize ETFs and their trading activity by their level of activeness, using the median value of their activeness index as the threshold. ETFs with activeness index values above the median are considered active, while those below the median are classified as passive. ETFs in the top quintile are deemed very active, while those in the bottom quintile are considered very passive. Table 10 reports the distribution of the ETF activeness index in the cross-section of ETFs for our full sample periods and several subsamples. Additionally, it shows the proportion of ETFs (number of funs) and the AUM, grouped by activeness level for the same sample periods.

The mean activeness index value in our sample ranges from 87% to 90%, indicating that the average ETF in our sample is highly active. Over 94% of the funds are classified as either moderately active or very active. In terms of AUM, approximately 70% are managed by active ETFs, showing that while passive ETFs are less numerous, they tend to be larger in size. For our broader ETF sample, we corroborate the findings of Easley et al. (2021), showing that most ETFs can be classified as active investment vehicles. Additionally, our results are consistent with those of Ben-David et al. (2023), who noted that the ETF industry has evolved with the emergence of niche, highly specialized products, such as sector, thematic, industry, and smart-beta ETFs.

[24] By comparing funds to the market portfolio rather than the specific index they track, the activeness index captures ETFs that either offer exposure to narrow segments (active-in-function) or significantly deviate from the market by stock-picking and alternative weighting schemes in order to generate alpha (active-in-form).

Table 11
Stock return volatility, excess return volatility, and ETFs activeness.
This table presents the results of Fama-MacBeth regressions of next quarter's (*t+1*) total return volatility (σ_{t+1}) and excess return volatility (σ_{t+1}^{exc}) on the squared stock price fragility of the current quarter (t) and a set of control variables. Stock price fragility is estimated based on mutual fund data only (G^{MF}) or ETF data only (G^{ETF}). 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). 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 firm's age, share turnover, and the lagged dependent variable (σ_t, σ_t^{exc}). ***, **, * denotes statistical significance at the 10%, 5%, and 1%. Standard errors are based on Newey–West adjustments. t-statistics are reported in parentheses. The sample covers Q1 2009 to Q4 2018.

	Total return volatility (σ_{t+1})			Excesss return volatility (σ_{t+1}^{exc})					
				1-Factor σ	3-Factor σ	4-Factor σ	1-Factor σ	3-Factor σ	4-Factor σ
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\sqrt{G}^{ETF(Active)}$	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.	38,513	38,513	36,016	38,513	38,513	38,513	38,513	38,513	38,513
Adj. R^2	0.013	0.026	0.471	0.014	0.012	0.011	0.029	0.026	0.025

To test whether more active ETFs play a significant role in propagating fragility by attracting short-term, speculative trades, we need to decompose the fragility measure to account for the separate influences of active and passive ETFs on future return volatility. 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 follows:

$$G_{it}^{ETF} = \left(\frac{1}{\theta_{it}}\right)^2 (W^{Act}\Omega^{Act}W^{Act} + W^{Pas}\Omega^{Pas}W^{Pas} + 2W^{Act}\Omega^{Act,Pas}W^{Pas}). \tag{12}$$

To identify active ETFs, we split our sample at the median activeness index value (50% threshold) since this cutoff most likely includes both active-in-form and active-in-function ETFs (Easley et al., 2021). We then replicate the main specifications for Tables 4 and 6, including the decomposition of G^{ETF} into active and passive components.

The results in Table 11 show that, consistent with our hypothesis, the active ETF component of G^{ETF} is the main driver behind its ability to forecast next quarter total return volatility and excess return volatility. Column (1) shows that most of the observed relationship between G^{ETF} and volatility stems from the active ETFs component. Column (2) presents the results of an examination of the same relationship if we include G^{MF}, while column (3) shows the results when we include the full set of control variables. When G^{MF} is added along with the control variables, it turns indistinguishable from zero. Interestingly, the coefficient on $G^{ETF(Passive)}$ turns negative and statistically significant. This evidence is consistent with the claims of Easley et al. (2021) and Ben-David et al. (2023) that passive broad-based ETFs attract investors seeking diversification at a lower cost rather than performance chasing. Thus, it is expected that the passive component of G^{ETF} is more likely to reflect investor preferences for low-cost diversification, making it less susceptible to capturing non-fundamentally driven trading decisions. In columns (4) - (9), we observe that our results hold after controlling for market, SMB, and HML exposures, though some coefficients are slightly smaller. Prior literature suggests that these results may be due to fragility also influencing the volatility of factor returns (Greenwood and Thesmar, 2011; Huang et al., 2021).

To summarize, this analysis suggests that the active ETF component of G^{ETF} primarily drives the observed relationship with future stock return volatility. Previous research shows that active ETFs often attract investors driven by sentiment, thematic investing, or short-term speculation (Israeli et al., 2017; Davies, 2022; Ben-David et al., 2023). Our results support these findings by showing that using ETF data significantly improves stock price fragility estimation. Additionally, we find that relying only on mutual fund data fails to adequately capture of non-fundamental demand shocks, leading to an underestimation of firms' exposure to this risk.

5. Conclusion

In this paper, we develop an ETF-based price fragility measure and compare it to the original measure developed by Greenwood and Thesmar (2011). We find that an ETF-based fragility measure (G^{ETF}) strongly predicts next-quarter stock return volatility. In comparison, the forecasting performance of the mutual fund-based measure (G^{MF}) significantly decreased out-of-sample and turns statistically indistinguishable from zero when we combine both measures in a regression. In addition, we analyze stock return comovement and find that the ETF-based measure is a stronger predictor of future daily return covariance (correlation).

Our evidence remains robust across various regression specifications and after including several control variables and the current quarter's dependent variable. Additionally, our findings are not limited to firms of a specific size or industry affiliation.

We also examine the implications of differences in ownership composition on our results. Our findings indicate that G^{ETF} partially captures the impact of institutional ownership on stock volatility, an effect that is absent when using G^{MF}. These results are a significant contribution of our study, as our methodology accounts for the rise of passive investing. In other words, our approach considers the increasing inclusion of ETFs in the portfolios of both retail and institutional investors.

While broad-based passive ETFs are widely known for providing diversification, we explore the role that more specialized and active ETFs may play in channeling non-fundamentally demand shocks and its relation to our previously documented evidence. Recent studies point to more active ETFs attracting the attention of investors whose trading decisions are likely driven extrapolation beliefs and sentiment (Huang et al., 2020; Easley et al., 2021; Ben-David et al., 2023). We find that our previously documented evidence stems from active ETFs.

Our results have significant implications for empirical asset pricing studies, particularly in the context of the impact of non-fundamentally driven demand shocks on stock return volatility and comovement. Although our approach does not completely eliminate the limitations associated with empirically estimating stock fragility, it offers a method that is not affected by many of the issues related to using mutual fund flows. This provides researchers with a more accurate proxy for assessing firm-level exposure to non-fundamental demand risk. Broadly, our findings contribute to the ongoing discussion about the repercussions of the rise in passive investments on market efficiency, a topic of great interest to investors, policymakers, and investment managers.

CRediT authorship contribution statement

Hamilton Galindo Gil: Writing – review & editing, Writing – original draft, Validation, Supervision, Investigation, Conceptualization. **Renato Lazo-Paz:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Theoretical model of stock price fragility

In this Appendix, we propose a theoretical model for stock price fragility. Specifically, we extend the model developed by Merton (1971) to consider an idiosyncratic liquidity shock in an economy with two agents who are heterogeneous in preferences. We begin by defining the agents' preferences, which 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 denotes the impatience rate, which is the same for both agents, and c_{it} denotes 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 dynamics of asset prices are exogenous and follow the equations:

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

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

where α denotes the expected rate of return of the risky asset. We assume that this asset does not have dividends as it is common for mutual funds to reinvest all profits in the portfolio. The volatility of risky asset returns is denoted by σ, and r denotes the risk-free interest rate. The aggregate shock in this economy is denoted by dZ_t, where Z_t denotes a standard Brownian motion. Additionally, the wealth dynamic of the agent i evolves according to Eq. (A.3):

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

where θ_i denotes 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 idiosyncratic shock is denoted 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 occurs when $dZ_{i,Liq}$ is negative, which means that the agent suddenly experiences a drop in wealth. Eq. (A.3) in compact form is given by:

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

where

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

$$q_i = \begin{bmatrix}\theta_i\sigma & \sigma_{i,Liq}\end{bmatrix} \tag{A.6}$$
$$d\widetilde{Z}_i = \begin{bmatrix}dZ_t & dZ_{i,Liq}\end{bmatrix}'. \tag{A.7}$$

We now define the consumption–portfolio choice problem for the agent i as:

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

subject to:

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

with the following constraint:

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

where W_{i0} is the initial wealth of agent i. The stochastic optimal control problem (Eqs. (A.8), (A.9), and (A.10)) can be transformed into a dynamic stochastic programming problem, represented by the Hamilton–Jacobi-Bellman equation as follows:

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

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

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

A.1. Non-fundamental demand of the risky asset

We now calculate the total demand for the shares of the risky asset, denoted as N^d:

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

where N_i denotes 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:

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

Then, we obtain the shares demand of agent i as follows:

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

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

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

which is the share demand of the risky asset. By ordering the elements of Eq. (A.16), we have:

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

Eq. (A.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 Itô's lemma, we find the dynamics of risky-shares demand, dN^d, given by:

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

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

$$dN^d = \underbrace{\frac{1}{P_t}g(W_{1t},W_{2t})dt + \frac{1}{P_t}h(W_{1t},W_{2t})dZ_t}_{=dN_f:\text{change in fundamental demand}} + \underbrace{\frac{1}{P_t}[\theta_{1t}W_{1t}\sigma_{1,Liq}]dZ_{1,Liq} + \frac{1}{P_t}[\theta_{2t}W_{2t}\sigma_{2,Liq}]dZ_{2,Liq}}_{=dN_f:\text{change in non-fundamental demand}} \tag{A.19}$$

Then, Eq. (A.19) could be expressed as:

$$dN^d = dN_f + dN_{nf}, \tag{A.20}$$

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

$$dN_{nf} = \frac{1}{P_t}[\theta_{1t} W_{1t} \sigma_{1,Liq}] dZ_{1,Liq} + \frac{1}{P_t}[\theta_{2t} W_{2t} \sigma_{2,Liq}] dZ_{2,Liq}. \tag{A.21}$$

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

$$\theta_{it} = \frac{P_t N_{it}}{W_{it}} \longrightarrow N_{it} = \frac{\theta_{it} W_{it}}{P_t}. \tag{A.22}$$

We introduce the expression $\theta_{it} W_{it}/P_t$ for both agents $i = 1, 2$ into Eq. (A.21), resulting in:

$$dN_{nf} = N_{1t}\sigma_{1,Liq} dZ_{1,Liq} + N_{2t}\sigma_{2,Liq} dZ_{2,Liq}. \tag{A.23}$$

This equation reflects the effects of liquidity shocks of two agents on 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. By dividing Eq. (A.23) by the total shares outstanding, N, and considering that η_{it} denotes the ownership of agent i in the risky asset at time t, where $\eta_{it} = N_{it}/N$, we have:

$$dN_{nf} = N\eta_{1t}\sigma_{1,Liq} dZ_{1,Liq} + N\eta_{2t}\sigma_{2,Liq} dZ_{2,Liq}. \tag{A.24}$$

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 if agent 2 experiences the same shock. The reason for this is that agent 1 has a higher concentration of the asset in his portfolio. Therefore, ownership is relevant to understanding the effects of liquidity shocks on asset demand and, consequently, on asset prices.

A.2. Stock price fragility

We define fragility as the expected volatility of non-fundamental demand given an asset's ownership structure, based on Greenwood and Thesmar (2011). In our theoretical model, shifts in non-fundamental demand are represented by Eq. (A.24). Although its expected value is zero, $E(dN_{nf}) = 0$, its variance aligns with the asset fragility definition of Greenwood and Thesmar (2011). Therefore, we define asset fragility as the variance of dN_{nf} as follows:

$$\text{Fragility} = Var(dN_{nf}). \tag{A.25}$$

To explicitly account for asset ownership and the variance–covariance matrix of liquidity shocks, we express dN_{nf} in matrix form as follows:

$$dN_{nf} = \underbrace{\begin{bmatrix} N\eta_1 & N\eta_2 \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} \sigma_{1,Liq} dZ_{1,Liq} \\ \sigma_{2,Liq} dZ_{2,Liq} \end{bmatrix}}_{Z} \equiv MZ. \tag{A.26}$$

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

$$\begin{aligned} Var(dN_{nf}) &= E\left[MZZ'M'\right], \quad \text{with } E[dN_{nf}] = 0 \\ &= ME\left[ZZ'\right]M' \\ &= N^2 \begin{bmatrix} \eta_1 & \eta_2 \end{bmatrix} \Omega \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}, \end{aligned} \tag{A.27}$$

where $\begin{bmatrix} \eta_1 & \eta_2 \end{bmatrix}$ is a vector of asset ownership and Ω denotes the Var-Cov matrix of liquidity shocks, defined as:

$$E\left[ZZ'\right] = \Omega = \begin{bmatrix} \sigma^2_{1,Liq} Var(dZ_{1,Liq}) & \sigma_{1,Liq}\sigma_{2,Liq} Cov(dZ_{1,Liq}, dZ_{2,Liq}) \\ \sigma_{1,Liq}\sigma_{2,Liq} Cov(dZ_{1,Liq}, dZ_{2,Liq}) & \sigma^2_{2,Liq} Var(dZ_{2,Liq}) \end{bmatrix}. \tag{A.28}$$

In our model, we assume that both idiosyncratic shocks are independent, so $Cov(dZ_{1,Liq}, dZ_{2,Liq}) = 0$. However, the model can easily be extended to the case where these shocks are correlated. Using Eq. (A.28), our definition of fragility would be:

$$\text{Fragility} = Var(dN_{nf}) = N^2 \begin{bmatrix} \eta_1 & \eta_2 \end{bmatrix} \Omega \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}, \tag{A.29}$$

which considers the effect of ownership and the variance–covariance matrix of liquidity shocks. This result provides a microfoundation for the measure of stock fragility defined by Greenwood and Thesmar (2011).

A.3. Stock price fragility and stock return volatility

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

$$N_t^s = AP_t, \quad A > 0. \tag{A.30}$$

In equilibrium, we have:

$$dN_t^s = dN_t^d. \tag{A.31}$$

Using the Eqs. (A.20) and (A.30), the equilibrium condition (A.31) is equivalent to:

$$d(AP_t) = dN_f + dN_{nf}. \tag{A.32}$$

By dividing by P_t and applying the variance operator in Eq. (A.32), we have:

$$\mathrm{Var}\left[\frac{dP_t}{P_t}\right] = \frac{1}{A^2P_t^2}\mathrm{Var}[dN_f] + \frac{1}{A^2P_t^2}\underbrace{\mathrm{Var}[dN_{nf}]}_{\text{Fragility}}, \tag{A.33}$$

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

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.finmar.2024.100946.

Data availability

Data will be made available on request.

References

Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. J. Financial Mark. 5 (1), 31–56.
Antón, M., Polk, C., 2014. Connected stocks. J. Finance 69 (3), 1099–1127.
Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. J. Finance 61 (4), 1645–1680.
Ben-David, I., Franzoni, F., Kim, B., Moussawi, R., 2023. Competition for attention in the ETF space. Rev. Financ. Stud. 36 (3), 987–1042.
Ben-David, I., Franzoni, F., Moussawi, R., 2017. Exchange-traded funds. Annu. Rev. Finan. Econ. 9 (1), 169–189.
Ben-David, I., Franzoni, F., Moussawi, R., 2018. Do ETFs increase volatility?: Do ETFs increase volatility? J. Finance 73 (6), 2471–2535.
Ben-David, I., Franzoni, F., Moussawi, R., Sedunov, J., 2021. The granular nature of large institutional investors. Manage. Sci. 67 (11), 6629–6659.
Berger, E.A., 2022. Selection bias in mutual fund fire sales. J. Financ. Quant. Anal. 58 (3), 1039–1077.
Berk, J.B., Green, R.C., 2004. Mutual fund flows and performance in rational markets. J. Polit. Econ. 112 (6), 1269–1295.
Broman, M.S., 2016. Liquidity, style investing and excess comovement of exchange-traded fund returns. J. Financial Mark. 30, 27–53.
Brown, D.C., Davies, S.W., Ringgenberg, M.C., 2021. ETF arbitrage, non-fundamental demand, and return predictability. Rev. Finance 25 (4), 937–972.
Bushee, B.J., 2001. Do institutional investors prefer near-term earnings over long-run value? Contemp. Account. Res. 18 (2), 207–246.
Bushee, B.J., Noe, C.F., 2000. Corporate disclosure practices, institutional investors, and stock return volatility. J. Account. Res. 38 (2000), 171–202.
Campbell, J.Y., Lettau, M., Malkiel, B.G., Xu, Y., 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. J. Finance 56 (1), 1–43.
Carhart, M.M., 1997. On persistence in mutual fund performance. J. Finance 52 (1), 57–82.
Coval, J., Stafford, E., 2007. Asset fire sales (and purchases) in equity markets. J. Financ. Econ. 86 (2), 479–512.
Da, Z., Huang, X., Jin, L.J., 2020. Extrapolative beliefs in the cross-section: what can we learn from the crowds? J. Financ. Econ..
Da, Z., Shive, S., 2018. Exchange traded funds and asset return correlations. Eur. Financial Manag. 24 (1), 136–168.
Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. J. Finance 52 (3), 1035–1058.
Dannhauser, C.D., Pontiff, J., 2019. Flow. SSRN Electron. J. 12–26.
Davies, S.W., 2022. Speculation sentiment. J. Financ. Quant. Anal. 57 (7), 2485–2515.
De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990. Noise trader risk in financial markets. J. Polit. Econ. 98 (4), 703–738.
Dong, M., Hirshleifer, D., Teoh, S.H., 2021. Misvaluation and corporate inventiveness. J. Financ. Quant. Anal. 56 (8), 2605–2633.
Doshi, H., Elkamhi, R., Simutin, M., 2015. Managerial activeness and mutual fund performance. Rev. Asset Pricing Stud. 5 (2), 156–184.
Dou, W.W., Kogan, L., Wu, W., 2022. Common fund flows: Flow hedging and factor pricing. SSRN Electron. J..
Easley, D., Michayluk, D., O'Hara, M., Putniņš, T.J., 2021. The active world of passive investing. Rev. Finance 25 (5), 1433–1471.
Edmans, A., Goldstein, I., Jiang, W., 2012. The real effects of financial markets: The impact of prices on takeovers. J. Finance 67 (3), 933–971.
Evans, R.B., 2010. Mutual fund incubation. J. 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. J. Bus. 38 (1), 34–105.
Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. J. Financ. Econ. 33 (1), 3–56.
Fama, E.F., French, K.R., 2010. Luck versus Skill in the cross-section of mutual fund returns. J. Finance 65 (5), 1915–1947.
Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. J. Polit. Econ. 81 (3), 607–636.
Fang, L., Jiang, H., Sun, Z., Yin, X., Zheng, L., 2023. Limits to diversification: Passive investing and market risk. AFA Annu. Meet..
Francis, B., Hasan, I., Shen, Y.V., Ye, P., 2021. Stock price fragility and the cost of bank loans. J. Empir. Financ. 63, 118–135.

Frazzini, A., Lamont, O.A., 2008. Dumb money: Mutual fund flows and the cross-section of stock returns. J. Financ. Econ. 24.
Friberg, R., Goldstein, I., Hankins, K.W., 2024. Corporate responses to stock price fragility. J. Financ. Econ. 153.
Gabaix, X., 2011. The granular origins of aggregate fluctuations. Econometrica 79 (3), 733–772.
Gârleanu, N., Panageas, S., 2015. Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing. journal of political economy 16.
Glosten, L., Nallareddy, S., Zou, Y., 2021. ETF activity and informational efficiency of underlying securities. Manage. Sci. 67 (1), 22–47.
Greenwood, R., Thesmar, D., 2011. Stock price fragility. J. Financ. Econ. 102 (3), 471–490.
Huang, S., O'Hara, M., Zhong, Z., 2020. Innovation and informed trading: Evidence from industry ETFs. Rev. Financ. Stud. 00, 37.
Huang, S., Ringgenberg, M.C., Zhang, Z., 2022. The information in asset fire sales. Manage. Sci. 69 (9), 4973–5693.
Huang, S., Song, Y., Xiang, H., 2021. Noise trading and asset pricing factors. Manage. Sci. Forthcoming.
Israeli, D., Lee, C.M., Sridharan, S.A., 2017. Is there a dark side to exchange traded funds? An information perspective. Rev. Account. Stud. 22 (3), 1048–1083.
Karmaziene, E., Sokolovski, V., 2022. Short-selling equity exchange traded funds and its effect on stock market liquidity. J. Financ. Quant. Anal. 57 (3), 923–956.
Koijen, R.S., Yogo, M., 2019. A demand system approach to asset pricing. J. Polit. Econ. 127 (4), 1475–1515.
Li, J., 2022. What drives the size and value factors? Rev. Asset Pricing Stud. 12 (4), 845–885.
Li, F.W., Zhu, Q., 2022. Short selling ETFs. Rev. Asset Pricing Stud. 12 (4), 960–998.
Lou, D., 2012. A flow-based explanation for return predictability. Rev. Financ. Stud. 25 (12), 3457–3489.
Madan, D.B., 2010. Stochastic processes in finance. Annu. Rev. Finan. Econ. 2 (1), 277–314.
Madhavan, A., 2014. Exchange-traded funds: An overview of institutions, trading, and impacts. Annu. Rev. Finan. Econ. 6 (1), 311–341.
Massa, M., Schumacher, D., Wang, Y., 2020. Repurchases for price impact: Evidence from fragile stocks. SSRN Electron. J..
Merton, R.C., 1971. Optimum consumption and portfolio rules in a continuous-time model. J. Econom. Theory 3 (4), 373–413.
Pavlova, A., Sikorskaya, T., 2023. Benchmarking intensity. Rev. Financ. Stud. 36 (3), 859–903.
Pohl, W., Schmedders, K., Wilms, O., 2021. Asset pricing with heterogeneous agents and long-run risk. J. Financ. Econ. 941–964.
Ross, S.A., 1976. The arbitrage theory of capital asset pricing. J. Econom. Theory 13 (3), 341–360.
Rzeznik, A., Weber, R., 2022. Money in the right hands: the price effects of specialized demand. SSRN Electron. J..
Sherrill, D.E., Shirley, S.E., Stark, J.R., 2017. Actively managed mutual funds holding passive investments: What do ETF positions tell us about mutual fund ability? J. Bank. Financ. 76, 48–64.
Sherrill, D.E., Shirley, S.E., Stark, J.R., 2020. ETF use among actively managed mutual fund portfolios. J. Financial Mark. 51 (1), 1–15.
Shleifer, A., Vishny, R.W., 1997. The limits of arbitrage. J. Finance 52 (1), 35–55.
Sias, R.W., 1996. Volatility and the institutional investor. Financ. Anal. J. 52 (2), 13–20.
Wang, J., 1994. A model of competitive stock trading volume. J. Polit. Econ. 102 (1), 127–168.
Wang, J., 1996. The term structure of interest rates in a pure exchange economy with heterogeneous investors. J. Financ. Econ. 41 (1), 75–110.
Wardlaw, M., 2020. Measuring mutual fund flow pressure as shock to stock returns. J. Finance 75 (6), 3221–3243.
Yan, X., Zhang, Z., 2009. Institutional investors and equity returns: Are short-term institutions better informed. Rev. Financ. Stud. 22 (2), 893–924.