A Bound on Price Impact and Disagreement*

Philippe van der Beck[†] Lorenzo Bretscher[‡] Julie Zhiyu Fu[§]

August 31, 2025

[Click here for the latest version]

Abstract

Asset prices are highly volatile, yet portfolio flows – changes in portfolio holdings – are relatively small. This reveals a fundamental tension between the price impact of portfolio flows and the agreement among investors: if price volatility is high while portfolio turnover is low, then either market participants largely agree with each other or they are not sensitive to price changes (they are "inelastic"), resulting in large price impacts of portfolio flows. We formalize this trade-off and demonstrate that the ratio of return volatility to portfolio turnover provides a lower bound on price impact, conditional on the level of investor disagreement. Using several measures from survey data, we document substantial disagreement, implying meaningful lower bounds on price impacts. The bounds align closely with reduced-form estimates from a variety of quasi-experiments, such as price impacts from index reconstitutions. We demonstrate how these bounds vary across horizons, different assets, and at various levels of aggregation, including the aggregate stock market, and discuss their implications for asset pricing models. We argue that in such markets with high disagreement and price impact, observed trading activity is not peripheral but central to understanding asset price movements.

*We thank Avinash Sattiraju for excellent research assistance. We thank Malcolm Baker, John Campbell, Xavier Gabaix, Robin Greenwood, Valentin Haddad, Sam Hanson, Paul Huebner (discussant), Ralph Koijen, Tyler Muir, Andrei Shleifer, Erik Stafford, Adi Sunderam, Luis Viceira, Motohiro Yogo, and seminar participants at the Demand in Asset Markets Workshop, the 11th BI-SHOF Conference on Asset Pricing and Financial Econometrics, Harvard Business School, the University of Hong Kong (HKU), and the Hong Kong University of Science and Technology (HKUST) for helpful comments and discussions.

[†]pvanderbeck@hbs.edu, Harvard Business School

[‡]lorenzo.bretscher@unil.ch, University of Lausanne, Swiss Finance Institute & CEPR

[§]z.fu@wustl.edu, Washington University in St. Louis, Olin School of Business

1 Introduction

It is uncontroversial that investors often disagree with each other. Such disagreement gives rise to heterogeneous portfolio holdings and flows. However, how much flows matter for asset prices is still an active debate, reflecting two different views on asset pricing. In one view, investors are highly elastic with respect to asset prices, and hence the market clears trading with little price adjustment. Under this view, quantity data – portfolio holdings and flows – tells us little about the drivers of asset prices. By contrast, in the inelastic market view, quantities have a significant impact on prices, making portfolio holdings and flows central to understanding asset prices.

In this paper, we argue that, given observable moments on quantities and prices, if we acknowledge that investors disagree with each other, we must accept the inelastic market view – that trading has a large price impact. Our argument rests on a simple observation: asset prices are highly volatile, yet portfolio flows – changes in portfolio holdings – are relatively small. This observation creates a trade-off between *investor agreement* and *price impact*: with large agreement among investors, one can easily generate little (or no) trading alongside volatile prices. However, with large disagreement, one might expect to observe large trading among investors. Yet we do not, despite high price volatility. Hence, this observation must imply that investors are insensitive to price changes, i.e., price-inelastic. As a result, small portfolio flows have large price impacts. This fundamental tension between price impact and agreement forms the core insight of our paper.

We formalize the trade-off between price impact and agreement mathematically as a simple bound based on two directly observable moments: portfolio flow volatility σ_q and return volatility σ_p. Within our framework, *demand shifts* are defined as any change in investors' portfolio choice, holding prices constant (e.g., cash flow news). We define *investor agreement* ρ, loosely speaking, as the correlation of demand shifts across investors.[1] Demand shifts can diverge for many reasons beyond disagreement about cash-flow news – for example, disagreement about discount rates, changes in regulatory frictions, etc. Importantly, what matters is the agreement in demand *shifts*, as opposed to agreement in *levels*. Investors may strongly disagree about the level of prices, yet still largely agree on how new information changes prices. *Price impact* $\mathcal{M}$ is given by the percentage change in prices per 1% demand shift.

Our main theoretical result establishes that the price impact $\mathcal{M}$ is bounded from below as follows:

$$\mathcal{M} \geq \frac{\sigma_p}{\sigma_q} \times \sqrt{\frac{1}{\rho} - 1}. \tag{1}$$

[1] Formally, ρ is defined as the share of demand shift variation explained by the size-weighted cross-investor average.

The bound captures our key insight: high return volatility σ_p relative to portfolio flow volatility σ_q (i.e., a high p/q volatility ratio) implies either a high price impact (high $\mathcal{M}$) or high agreement among investors (high ρ). To understand the intuition of our bounds, consider the analogy of total demand shifts as an iceberg. The size of an iceberg is determined by the total of demand shifts. Only disagreement surfaces and becomes visible as observable portfolio flows, while the common demand shifts that move all investors in the same direction remain submerged and unobserved. Investor agreement ρ determines the relative size of the observable flows versus the unobserved common demand shifts. When investors perfectly agree with each other in updating their beliefs ($\rho \to 1$), observed portfolio flows represent only small tip of a large iceberg. Price variation is primarily driven by large unobserved common demand shifts. In this case, the high p/q volatility ratio are uninformative about the price impact. Conversely, when investors disagree with each other (small ρ), the visible flows represent a significant portion of the total demand shifts. Hence, low observed portfolio flows imply small total demand shifts, and large return volatility can only be reconciled with a substantial price impact.

These two scenarios, while both consistent with observed flow and return volatilities, have fundamentally different implications for our understanding of asset prices. In the case of elastic investors that largely agree with each other, portfolio holdings and flows are merely a sideshow – they capture at best minor deviations from common demand shifts which drive prices. In such a world, asset prices can be well-described by representative agent models, while portfolio flows across investors are a largely irrelevant byproduct of price formation. In contrast, in the case of inelastic investors that disagree with each other, portfolio flows carry substantial incremental information about asset prices. Here, observed quantities are no longer irrelevant but central to our understanding of asset prices.

The derivation of our bound relies on a set of fairly general assumptions. Most notably, following the long tradition of log-linearization in finance and the growing literature on demand-system asset pricing (Campbell and Viceira, 2002; Koijen and Yogo, 2019), we assume that portfolio choice problems can be approximated to a first order by linear demand curves. Importantly, we impose no structural assumptions on the source or structure of demand shifts, nor on particular microfoundations of the elasticities. This makes our bound *empirical* in nature, similar in spirit to the Hansen–Jagannathan bound (Hansen and Jagannathan, 1991). Due to the model-agnostic nature of the bound, it can serve as a diagnostic tool when developing structural models to rationalize these moments. It can also be used in empirical studies when estimating price impact $\mathcal{M}$ or agreement ρ to back out the other parameter, thus providing a more comprehensive picture of the market environment.

We apply our bound in Equation (1) to the US stock market. Returns are volatile: an average stock has a return volatility of $\sigma_p \approx 20\%$ at the quarterly frequency. On the other hand, flow volatility is not much larger. We show that flow volatility σ_q can be simply measured with "portfolio turnover", the average net portfolio reallocations across all investors over the period of interest. At the quarterly frequency, the average stock has a portfolio turnover of around 20%, resulting in a p/q volatility ratio of $\frac{\sigma_p}{\sigma_q} \approx 1$. The common assertion of excess trading activity is typically based on the gross trading volume, which is the sum of all trading in one period, including transitory round-trip trades. It significantly overstates the actual portfolio turnover that persists over the period of interest. This parallels to how return volatility is measured – using net price changes over given periods, rather than summing all instantaneous price changes within periods. Similarly, we should also net out intra-period round-trip trades when assessing flow volatility.

Flow and return volatilities are readily measurable in the data but investor agreement ρ, defined as the correlation of investors' demand shifts, is inherently unobservable. However, there is substantial empirical evidence confirming that investors' demand shifts are far from perfectly correlated ($\rho \ll 1$). Investors differ markedly in their cash-flow forecasts, regulatory constraints, return expectations, trading patterns, and portfolio compositions.[2] Given this evidence, a highly elastic market at the quarterly frequency appears unlikely through the lens of our bound. For example, as $\frac{\sigma_p}{\sigma_q} \approx 1$ for the average stock, achieving a price impact below 0.1 requires almost perfect agreement among investors, i.e., $\rho > 0.99$.

To estimate the price impact bound, we incorporate empirical proxies for investor agreement ρ. Our goal is not to obtain precise point estimates, but to demonstrate that ρ lies within a moderate range – avoiding pathological extremes near zero or one. Our baseline proxy measures agreement through the common variations in forecast updates on earnings across analysts. For the average U.S. firm, forecast updates of earnings per share (EPS) across analysts explain approximately 58% of the total variation in one-quarter-ahead EPS updates, yielding $\rho = 0.58$. When we use long-term growth (LTG) forecasts instead, we obtain $\rho = 0.27$. Applying these agreement measures to our model generates stock-level price impacts of 0.75 and 1.0, respectively. We also draw on proxies for ρ from event studies and structurally inferred demand shifts, and find no evidence that ρ takes implausibly high values.

The model parameter ρ should reflect all forms of heterogeneity in demand shifts, including belief

[2]See, among others, Giglio et al. (2021), Koijen and Yogo (2019), Dahlquist and Ibert (2024), Couts et al. (2024), Barber and Odean (2008), Guiso et al. (2008), Kandel and Pearson (1995), Barber and Odean (2001), Bretscher et al. (2025), and Bretscher et al. (2025). In addition, the vast investor disagreement can be directly observed by the fact that the number of different mutual funds catering to the preferences and beliefs of different investors exceeds the number of stocks in the U.S. (see Investment Company Institute (2025)).

disagreement, regulatory constraints, and preferences, which may not be fully reflected in our empirical proxies for ρ. However, our measures suggest that the true value of ρ lies in the moderate range rather than at the extremes. In the moderate range, $\sqrt{\frac{1}{\rho} - 1}$ is relatively flat, making it insensitive to variations in ρ. Consequently, cross-stock variation in our bounds is driven primarily by the p/q volatility ratio, $\frac{\sigma_p}{\sigma_q}$. In fact, a simplified price impact estimate $\tilde{\mathcal{M}} \equiv \frac{\sigma_p}{\sigma_q}$ (which implicitly assumes $\rho = 0.5$) performs nearly as well as the general bound.

We empirically validate our price impact bounds against well-documented demand shock events, including S&P 500 index inclusions and mutual fund flow-induced trading. Our bound-implied price impact estimates exhibit strong correlations with actual price movements across these events. Stocks with larger bounds experience significantly higher price changes for a given demand shock. For instance, the price impact of flow-induced trades increases monotonically with our bounds. In contrast, traditional liquidity measures based on gross trading volume, such as Amihud's (2002) illiquidity ratio, show no significant explanatory power for price impacts of persistent demand shifts. Constructing our bounds with gross trading volume (rather than portfolio turnover) reveals no significant relationship to event study price impacts, highlighting that portfolio turnover is not merely a modeling choice but an economically meaningful quantity.

We further examine how the price impact bounds vary at different horizons, across assets, and at different levels of aggregation. First, price impact declines monotonically with horizon – that is, demand elasticities increase at longer horizons. Daily price impacts are substantially larger than quarterly impacts, which in turn exceed annual impacts. Second, while daily price impact has declined significantly from 1990 to 2024, quarterly and annual price impacts have remained largely unchanged. Third, consistent with information-based theories, large-cap stocks exhibit *smaller* price impact, while stocks with higher systematic risk show *larger* price impact, consistent with risk-based foundations. Momentum stocks also experience higher price impact, reflecting upward-sloping demand curves of momentum investors. Fourth, examining our bounds at different levels of aggregation reveals that portfolio turnover falls quickly as aggregation increases. This is intuitive as porfolio flows in similar assets offset each other.[3] Return volatility also declines with aggregation due to diversification of idiosyncratic risks. Overall, we find that turnover decreases at a higher rate compared to return volatility which implies that our price impact bounds rise monotonically with aggregation – from individual stocks, to industries, to size/value portfolios, and ultimately to the market, which shows the highest impact (Gabaix and Koijen (2021)).

[3]For example, trading Apple against Microsoft creates stock-level turnover but none at the tech-sector level.

Our bounds provide a measure of how informative portfolio flow data is for asset prices. When demand is highly elastic, portfolio flows reveal little about prices as demand shifts can be easily accommodated with minimal price impact. Conversely, when demand is inelastic, portfolio flows are highly informative for prices. Empirically, our bounds suggest that the stock market is closer to the "inelastic and heterogeneous" paradigm than to the "elastic and homogeneous" paradigm. In such markets, observed portfolio flows reveal a significant portion of the underlying demand variation, and have significant impact on prices. Echoing Hong and Stein (2007) and more recently Gabaix and Koijen (2021), our bounds demonstrate that trading volume is not a mere byproduct of price formation, but is essential for understanding asset prices and financial market volatility.

Related Literature. This paper offers a synthesis of the literatures on price impact and on investor disagreement. First, a large strand of the literature documents that investor-specific demand shifts can have a meaningful long-term price impact. For example, a series of papers studies price changes upon inclusion or deletion of a stock in an index (see, among others, Shleifer, 1986, Harris and Gurel (1986), Wurgler and Zhuravskaya (2002), Kaul et al. (2000), Chang et al. (2015), Pavlova and Sikorskaya (2022), Greenwood and Sammon (2025), and Aghaee (2025)). Coval and Stafford (2007), Lou (2012), and Edmans et al. (2012) document persistent price changes due to flow-induced trading by mutual funds. Hartzmark and Solomon (2021), Schmickler and Tremacoldi-Rossi (2022), Kvamvold and Lindset (2018), and Honkanen et al. (2025) document that reinvested dividend payouts significantly affect asset prices. Our price impact bounds provide an ex ante statistic that serves as a simple benchmark based on observable empirical moments. In addition, finding exogenous variation in demand to identify price impact for aggregated portfolios such as the total stock market is often difficult. Our bounds can easily be computed at different levels of aggregation for all asset classes and thus provide a useful sanity check for event studies. More importantly, event-study evidence is often difficult to obtain, particularly when researchers face limited cross-sectional variation over time or insufficient time-series variation across assets. Estimating price impacts for aggregated portfolios – such as entire asset classes – is especially challenging, as it requires identifying demand shifts that are both exogenous and sufficiently large. Our estimation-free bounds offer a practical alternative by providing theoretically grounded benchmarks for the expected price impact in settings where event studies are infeasible.

More broadly, our bounds contribute to the burgeoning demand-system asset-pricing literature that jointly models investors' portfolio allocation and asset prices (Koijen and Yogo, 2019; Gabaix and

Koijen, 2021).[4] Several recent papers in this literature have alluded to the tension between elasticity and investor heterogeneity. Gabaix and Koijen (2021) argue that the relatively stable equity share of institutional investors implies inelasticity at the aggregate market level. Their granular instrument variable (GIV) estimator imposes a factor structure on demand shifts and identifies investor agreement by extracting common factors from investor flows. More recently, Gabaix et al. (2025) compute risk transfer – changes in market risk exposure by households, a measure conceptually close to the portfolio turnover for the aggregate market – is very small at the quarterly frequency. They demonstrate that standard macro-finance models with high price elasticities cannot reconcile the tension between the heterogeneity in holdings and the small risk transfer in flows. Complementary to their approach, our bounds are effectively *model-free*. We do not take a stand on a specific model linking heterogeneous portfolios to unobserved demand shifts. Instead, we use observed flow and price volatility to bound investor disagreement and price impact. Moreover, we compute bounds for individual stocks, different portfolios, and the aggregate stock market, and test their empirical relevance in event studies.

Our paper also contributes to the literature on investor disagreement. For example, Kandel and Pearson (1995) and Bamber et al. (1999) document that earnings announcement days consistently feature abnormally high trading volume and small price changes. In those papers, the combination of high volume and low volatility is typically interpreted as evidence of differential interpretations of public signals, i.e., *disagreement*. Hong and Stein (2007) advocate for models featuring disagreement, given the enormous trading volume observed even at times when return volatility is low.[5] While we find that portfolio turnover is low at longer horizons, we argue that this does not reflect an absence of disagreement. Instead, it reflects the inelasticity of market participants, which amplifies the price impact of investor-specific demand shifts and serves as an empirically useful measure of long-term price impact.

Third, by drawing a distinction between gross trading volume and portfolio turnover, our paper naturally contributes to the literature on market liquidity.[6] We highlight that the distinction between trading volume and portfolio turnover helps explain why traditional liquidity measures – such as those in Amihud (2002), Pástor and Stambaugh (2003), and Brennan et al. (2013) – may be less suitable

[4]See, among others, Koijen and Yogo (2020), Haddad et al. (2021), Han et al. (2021), Koijen et al. (2021), Fang et al. (2022), Coqueret (2022), Huebner (2023), Jiang et al. (2022), Jiang et al. (2024), Koijen et al. (2024), Jansen (2025), Tamoni et al. (2024), Bretscher et al. (2025), Chaudhary et al. (2024), Jansen et al. (2024), Darmouni et al. (2022).

[5]See, for example, Harris and Raviv (1993) and Banerjee and Kremer (2010) for theoretical models reconciling the observed empirical patterns.

[6]See, among others, Constantinides (1986), Brennan and Subrahmanyam (1996), Heaton and Lucas (1996), Vayanos (1998), Brennan et al. (1998), Datar et al. (1998), Chordia et al. (2001), Amihud (2002), Jones (2002), Huang (2003), Pástor and Stambaugh (2003), Anshuman and Viswanathan (2005), Brunnermeier and Pedersen (2009), Garleanu and Pedersen (2007), and Bouchaud (2022).

for capturing long-term price impact from persistent demand shifts. Intuitively, our bounds on price impact can be viewed as a long-horizon counterpart to the illiquidity measure proposed by Amihud (2002).

The remainder of the paper is structured as follows. Section 2 lays out the main theory. Section 3 describes the data, construction of portfolio turnover and discusses its differences from gross trading volume. Section 4 constructs the price impact bounds for the cross-section of U.S. equities; Section 5 tests the empirical relevance of the bounds using different event studies. Motivated by the empirical relevance, Section 6 then examines the heterogeneity of these bounds outside of event studies – over time, across assets, and at different levels of aggregation. Section 7 concludes.

2 Theory

In this section, we lay out our main framework and derive the price impact bound.

Notation. Throughout, we use $i = 1, 2, ..., I$ to denote the investor, n to denote the asset. We use $S_i(n)$ to denote the ownership share of investor i in the market for asset n. As a short-hand, we use subscript S in place of i to denote size-weighted aggregation, i.e., $x_S(n) = \sum_{i=1}^{I} S_i(n)x_i(n)$. To highlight the cross-sectional expectation-like feature of the size-weighted aggregation, we also use $\hat{\mathbb{E}}_S^{cs}[x_i] = \sum_{i=1}^{I} S_i x_i$, and suppress the subscript S when there is no ambiguity.

2.1 The Demand Curve

For illustration purposes, we start by deriving the price impact bound in a single-asset portfolio-choice model. Then we show in Section 2.6 that our bound applies to a multi-asset setting as well.

Consider a generic portfolio allocation $Q_{i,t}(n) = Q_i(P_t(n), U_{i,t})$, where $Q_{i,t}(n)$ is the quantity of asset n held by investor i at time t, $P_t(n)$ is the price of asset n at time t, and $U_{i,t}$ captures all other factors that affect investor i's demand for asset n at the given price $P_t(n)$. These factors can include the investor's wealth, the risk-free rate, risk aversion, uncertainty, prices of substitutable assets, and other relevant variables.

We take a log-linear approximation of the portfolio choice problem around the long-run mean and take first-differences to obtain a linear demand curve:

$$\Delta q_{i,t}(n) = -\zeta_i(n)\Delta p_t(n) + u_{i,t}(n) \tag{2}$$

where $\Delta q_{i,t}(n)$ is the percentage change in holdings of asset n by investor i at time t (which we refer to as *portfolio flows* or simply *flows*), $\Delta p_t(n)$ is the percentage price change of asset n at time t, referred to as its return at time t, and $u_{i,t}(n)$ is the demand shift for investor i at time t. For simplicity, we assume that their time-series means are equal to zero. The parameter $\zeta_i(n)$ is the investor-asset-specific elasticity, which measures how much investor i's demand for asset n changes when the price changes by 1%, *ceteris paribus*.

To provide intuition for the different components of the demand curve, we can connect this linear specification with canonical models. In Appendix B, we sketch several microfoundations, including standard portfolio choice under CRRA utility. Our preferred interpretation draws on learning-from-price models such as Grossman and Stiglitz (1980) and Hellwig (1980). In these models, the demand shift $U_{i,t}$ represents noisy private signals about the asset's fundamental value, while the price $P_t(n)$ aggregates information across investors. The price elasticity $\zeta_i(n)$ captures the trade-off between the informativeness of one's private signal and the market price: the more accurate the market price is relative to the investor's private signal, the more the investor relies on the price, resulting in a more inelastic demand curve (smaller $\zeta_i(n)$).

While learning-from-price models provide a natural framework for interpreting elasticity and disagreement, we do not restrict our analysis to this interpretation. Instead, we specify the demand curve generically. In any asset pricing model that features portfolio choice, either explicitly or implicitly, investor demand can be decomposed into changes due to price movements and changes holding prices fixed.[7] Our bound holds under these different model frameworks. Moreover, the underlying model does not need to be static: in dynamic settings, investors care about the path of future expected returns, while the market clears through the current price, which summarizes the market's expectations about future returns. In this case, investors' demand shifts contain beliefs about future expected returns that deviate from those implied by the current price. Furthermore, our framework also readily accommodates multiple assets – as shown in Section 2.6, a multi-asset system also yields a single-price demand representation as in (2). Our only assumption at this stage is that a first-order log-linearization provides a reasonable approximation of the true portfolio choice problem.

For a generic demand curve specified in Equation (2), our goal is to connect elasticity and the correlation of demand shifts (agreement) to observable moments: volatilities in returns and portfolio flows. To do so, we first study how the return and flow volatilities are determined in the log-linear model.

[7] See Koijen and Yogo (2025) for further discussion on microfoundations.

In the remainder of this section, we proceed stock by stock and suppress the stock index n for notational ease, re-introducing it in the multi-asset extension.

2.2 Elasticity and Price Impact

The flip side of elasticity is the price impact per unit of demand shift. To see that, we impose the market clearing condition – all trades sum to zero. Denote $S_i = \frac{Q_i}{\sum_i Q_i}$ the ownership share of investor i in the market for asset n. The market clearing condition is given by:

$$\sum_i S_i \Delta q_{i,t} = 0 \tag{3}$$

Price adjusts to clear the market, and hence,

$$\Delta p_t = \frac{1}{\zeta_S} u_{S,t} \tag{4}$$

where $\zeta_S = \sum_i S_i \zeta_i$ and $u_{S,t} = \sum_i S_i u_{i,t}$ are the aggregate elasticity and demand shift given by the size-weighted averages of investor-specific elasticities and investor-specific demand shifts respectively. The inverse of the aggregate elasticity, $\frac{1}{\zeta_S}$, quantifies how much the price adjusts for a 1% aggregate demand shift of total outstanding shares. Therefore, the lower the aggregate demand elasticity, the larger is the price adjustment per unit of demand shift which is needed to induce investors to step in. We denote the inverse of the aggregate demand elasticity as $\mathcal{M}$ and refer to it as the *price impact* or *multiplier* of asset n.

The price impact $\mathcal{M}$ links return volatility to the volatility of the aggregate demand shift, given by:

$$\sigma_p^2 = \mathcal{M}^2 \cdot \mathrm{Var}(u_{S,t}) \tag{5}$$

Through the lens of this framework, the well-known excess volatility puzzle implies that either standard models do not generate sufficiently volatile aggregate demand shifts, or that agents are too responsive to price changes in these models, i.e., the price impact $\mathcal{M}$ is too small.

2.3 Portfolio Flows and Investor Agreement

To illustrate the relationship between portfolio flows and investor agreement, we first consider the case with homogeneous elasticities across investors: $\zeta_i = \zeta_S = \zeta$. This assumption will be relaxed later. Under the homogeneous elasticity assumption, we can plug the equilibrium price equation (4) into the demand equation (2) to have:

$$\Delta q_{i,t} = u_{i,t} - u_{S,t}. \tag{6}$$

Hence, trades reflect the *differences* of investors' demand shifts from the average demand shift in the market.

The size-weighted average variance of $\Delta q_{i,t}$ is given by:

$$\begin{aligned} \sigma_q^2 &\equiv \sum_{i=1}^{I} S_i \operatorname{Var}(\Delta q_{i,t}) \\ &= \left(\sum_{i=1}^{I} S_i \operatorname{Var}(u_{i,t}) \right) - \operatorname{Var}(u_{S,t}) \end{aligned} \tag{7}$$

To derive the second equality, we use Equation (6) and the identity $\sum_{i=1}^{I} S_i \operatorname{Cov}(u_{i,t}, u_{S,t}) = \operatorname{Var}(u_{S,t})$. Hereafter, we refer to σ_q as *flow volatility*. It measures the total amount of trading activity by investors. The theoretical analysis focuses on flow volatility defined in (7); later we show that portfolio turnover across all investors is a close proxy for flow volatility, and use the terminology interchangeably when the distinction is unimportant.

Defining $\rho \equiv \frac{\operatorname{Var}(\sum_{i=1}^{I} S_i u_{i,t})}{\sum_{i=1}^{I} S_i \operatorname{Var}(u_{i,t})}$, we can rewrite flow volatility as follows:

$$\sigma_q^2 = \operatorname{Var}(u_{S,t}) \left(\frac{1}{\rho} - 1 \right). \tag{8}$$

We refer to ρ as *investor agreement* and $\mathcal{D} := \sqrt{\frac{1}{\rho} - 1}$ as *investor disagreement.* To understand the interpretation, note that it is the share of demand shifts that is explained by the size-weighted cross-sectional average of the demand shifts. To see this most clearly, we can use the cross-sectional expectation notation $\hat{\mathbb{E}}_S^{cs}$ to express it as follows:

$$\rho = \frac{\operatorname{Var}\left(\hat{\mathbb{E}}^{cs}\left[u_{i,t} \mid t\right]\right)}{\hat{\mathbb{E}}^{cs}\left[\operatorname{Var}\left(u_{i,t} \mid i\right)\right]} \in [0, 1]. \tag{9}$$

Empirically, ρ is the R^2 of the (size-weighted) time fixed effects of the demand shifts. As an R^2, it ranges between 0 and 1. When $\rho = 1$, all investors have identical demand shifts, and hence are homogeneous; when $\rho \to 0$, the demand shifts are completely heterogeneous across investors.[8] Alternatively, with

[8]With finite number of investors, $\rho \geq \frac{\sum_i S_i^2 \sigma_i^2}{\sum_i S_i \sigma_i^2}$ if the covariances of the demand shifts across investors are non-negative, and it reaches the lower bound when shocks are completely uncorrelated. It reaches zero only if investors demand completely offset each other in aggregate.

homoskedasticity, the agreement ρ can also be interpreted as the average pairwise correlation of the demand shifts $\rho \approx \sum_{i=1}^{I} \sum_{j \neq i} S_i S_j \text{ corr}(u_{i,t}, u_{j,t})$.[9]

Note that disagreement comes not only from differences in idiosyncratic demand shifts, but also from differences in the responses of different investors to common factors. To see this, suppose investor-specific demand shifts are determined by their differential exposure λ_i to a single common shock η_t, which has unitary variance, i.e., $u_{i,t} = \lambda_i \eta_t$. Let $\hat{\mathbb{E}}^{cs}$ denote the size-weighted cross-sectional average, we have:

$$\rho = \frac{\hat{\mathbb{E}}^{cs}\left[\lambda_i\right]^2}{\hat{\mathbb{E}}^{cs}\left[\lambda_i^2\right]} = \left(1 + \frac{\widehat{\text{Var}}^{cs}\left(\lambda_i\right)}{\hat{\mathbb{E}}^{cs}\left[\lambda_i\right]^2}\right)^{-1}. \tag{10}$$

This implies that investor agreement decreases in the variation of the exposures to the common shock η_t across investors. Further, agreement can be arbitrarily close to zero when the variation in λ_i relative to its mean is large. In sum, investors can have low agreement even if their idiosyncratic demand shifts can be fully explained by a common shock, η_t, provided their exposures to that shock differ.

With this interpretation, Equation (8) states that flow volatility is the product of the size of aggregate demand shifts and the amount of investor disagreement. The extent to which highly volatile aggregate demand translates into flow volatility (turnover) is driven by how dissimilar investors are in their demand shifts.

2.4 The Price Impact Bounds

To derive the bound, the key observation is that both price volatility (5) and flow volatility (8) depend on the volatility of the average demand shift in the market, but with a different coefficient: the multiplier $\mathcal{M}$ for return volatility and $\frac{1}{\rho} - 1$ for flow volatility. Taking the ratio of flow volatility in Equation (8) and return volatility in Equation (5), we have:

$$\mathcal{M} = \frac{\sigma_p}{\sigma_q} \times \underbrace{\sqrt{\frac{1}{\rho} - 1}}_{\mathcal{D}} \tag{11}$$

Equation (11) connects observable market quantities – price and flow volatilities – to the underlying elasticity and investor agreement. When prices exhibit high volatility relative to trading activity (a large $\frac{\sigma_p}{\sigma_q}$ ratio), two possible explanations emerge: either the price multiplier $\mathcal{M}$ is large, amplifying

[9] We can write $\text{Var}(u_{S,t}) = \sum_{i=1}^{I} S_i^2 \sigma_i^2 + \sum_{i=1}^{I} \sum_{j \neq i} S_i S_j \sigma_i \sigma_j \text{ corr}(u_{i,t}, u_{j,t})$, under homoskedasticity, $\sigma_i = \sigma_j = \sigma$, so we have $\rho = \sum_{i=1}^{I} S_i^2 + \sum_{i=1}^{I} \sum_{j \neq i} S_i S_j \text{ corr}(u_{i,t}, u_{j,t})$. The first term is the Herfindahl–Hirschman Index (HHI) of the ownership distribution, which vanishes to zero as N is large.

price responses to demand shifts, or investors strongly agree with each other on demand shifts ($\rho \to 1$), causing observed trading activity (the tip of the iceberg of total demand shifts) to significantly underrepresent the magnitude of underlying demand shifts.

So far, price impact $\mathcal{M}$ is derived under the homogeneous-elasticity assumption. To consider the case with heterogeneous elasticities, we make an assumption on the distribution of elasticities. Without getting too attached to a particular data-generating process, we consider the following environment:

Assumption 1. *The elasticity ζ_i for each investor i is drawn independently from the parameters governing the demand shift process $u_{i,t}$.*

Assumption 1 serves as a neutral benchmark, but it is not necessary for the main result. With an arbitrary data generating process of elasticities and the demand shifts, one can end up in the pathological case where the investors that receive larger demand shifts end up selling because they also react more to the price changes. In Appendix A, we discuss the more precise condition under which our main theorem holds.

Under Assumption 1, the implied multiplier will be even larger for a given level of investor disagreement $\mathcal{D}$. Intuitively, this is because heterogeneous elasticities induce trading for reasons other than disagreement on demand shifts. For example, consider the case where investors experience identical demand shifts. The price adjusts, but as investors respond to the price adjustment differently, they also want to trade with each other.

Hence, when an econometrician infers the magnitude of aggregate demand shifts from observed flow volatility under the assumption of homogeneous elasticities, this assumption leads to an *overestimation* of the underlying demand shifts. The overestimation occurs because some observed trading activity stems not from heterogeneous demand shifts but from investors' heterogeneous responses to price changes. Since the true aggregate demand shifts are smaller under heterogeneous elasticities, the actual price impact exceeds that given by (11). Formally, we establish the following theorem:

Theorem 1. *Under Assumption 1, the price impact $\mathcal{M}$ of demand shifts is lower-bounded by the p/q volatility ratio $\frac{\sigma_p}{\sigma_q}$, adjusted by investor disagreement $\mathcal{D} := \sqrt{\frac{1}{\rho} - 1}$:*

$$\mathcal{M} \geq \frac{\sigma_p}{\sigma_q} \times \sqrt{\frac{1}{\rho} - 1} \tag{12}$$

Proof. See Appendix A. □

In-sample bounds. Notice that although we express the bound in terms of population parameters, the identities used in deriving the bound all hold in sample as well. Hence we can express the bound using sample moments, given as:

$$\mathcal{M} \geq \frac{\hat{\sigma}_p}{\hat{\sigma}_q} \times \sqrt{\frac{1}{\hat{\rho}} - 1} \tag{13}$$

where $\hat{\sigma}_p$ and $\hat{\sigma}_q$ are the sample counterparts of price and flow volatilities, and $\hat{\rho}$ is the investor agreement of demand shifts within the sample period.

Moreover, the bound can be applied period by period, under the assumption that $\Delta q_{i,t}$ and Δp_t have mean zero in a given period t (which can be achieved by demeaning across t, assuming that means are stable):

$$\mathcal{M}_t \geq \frac{|\Delta p_t|}{\sqrt{\sum_{i=1}^{I} S_i \Delta q_{i,t}^2}} \times \sqrt{\frac{1}{\rho_t} - 1} \tag{14}$$

where $\rho_t \equiv \frac{(\sum_{i=1}^{I} S_i u_{i,t})^2}{\sum_{i=1}^{I} S_i u_{i,t}^2}$ is the investor agreement *in period t*.

2.5 Flow Volatility and Portfolio Turnover

The key input to our bound, flow volatility σ_q, is defined as the size-weighted average of investor-specific flow volatilities. Seemingly complicated, we show that it is closely related to the total trading activity from changes in investors' portfolios, which we term *portfolio turnover*. For a stock in a given quarter t, portfolio turnover is defined as the sum of the absolute values of quarter-on-quarter changes in positions of all investors, normalized by shares outstanding:

$$\text{Turnover}_t = \frac{\sum_i |\Delta Q_{i,t}|}{\bar{Q}} \tag{15}$$

where $\Delta Q_{i,t} = Q_{i,t} - Q_{i,t-1}$ is the change in position of investor i from $t-1$ to t, and $\bar{Q}$ is total supply.

Portfolio turnover measures the (size-weighted) *mean absolute deviation* (MAD) of flows:

$$\mathbb{E}\big[\text{Turnover}_t\big] = \mathbb{E}\left[\sum_i S_i \frac{|\Delta Q_{i,t}|}{S_i \bar{Q}}\right] = \sum_i S_i \mathbb{E}\left[|\Delta q_{i,t}|\right]. \tag{16}$$

It mirrors the definition of flow volatility, $\sigma_q \equiv \sqrt{\sum_{i=1}^{I} S_i \mathbb{E}\left[(\Delta q_{i,t})^2\right]}$, but with an $\mathcal{L}_1$-norm rather than an $\mathcal{L}_2$-norm. For common distributions, the mean absolute deviation $\mathbb{E}\left[|\Delta q_{i,t}|\right]$ is proportional to the standard deviation $\sigma_{q,i}$ by a constant factor ν determined by the underlying distribution. For example, for normally distributed $\Delta q_{i,t}$, $\nu = \sqrt{\frac{\pi}{2}} \approx 1.25$. Empirically, Appendix Figure E.1 shows that portfolio turnover scaled by $\sqrt{\frac{\pi}{2}}$ and σ_q are effectively equivalent with a cross-sectional correlation of around 0.9 and an OLS slope coefficient of 1.1. For this reason, we view the scaled portfolio turnover, $\sqrt{\frac{\pi}{2}}\text{Turnover}_t$, as an alternative (and more robust) estimator for σ_q, and use portfolio turnover to refer to σ_q at the conceptual level.

Portfolio turnover is closely linked to gross trading volume by construction. However, unlike gross trading volume, which aggregates *all* trades within a quarter, portfolio turnover omits offsetting round-trip trades and measures net quarter-on-quarter changes in portfolio holdings. To see this, consider an investor that moves from 1000 shares at t to 1100 shares at $t + \frac{1}{2}$, back to 1000 shares at $t + 1$. While the investor's gross volume is 200 shares, their portfolio turnover from t to $t + 1$ is $|\Delta Q_{i,t}| = 0$ shares. To understand the liquidity provision at the quarterly frequency (here t to $t + 1$), the intra-quarter round trips are irrelevant and hence netted out from portfolio turnover. The empirical application provides more details on the distinction between portfolio turnover and gross trading volume.

2.6 The Bound in the Multi-asset System

Our bound thus far applies to single assets, but real-world portfolio choice involves substitution across multiple assets. The existing literature emphasizes how ignoring cross-asset substitution biases price impact estimates: Chaudhary et al. (2023) demonstrates that neglecting heterogeneous substitution patterns in bond markets leads to biased cross-sectional estimates, while Haddad et al. (2025) show that identifying aggregate elasticities requires time-series variation even when accounting for heterogeneous substitution patterns.

Since our bound relies on time-series variation, it avoids the cross-sectional bias identified by Chaudhary et al. (2023). The bound remains valid—single-asset flow and return volatilities still capture the fundamental trade-off between price impact and investor agreement. However, both price impact and investor agreement acquire richer interpretations in multi-asset settings, which we now explore.

The critical insight is that multi-asset demand systems can still be expressed as single-price demand equations of the form in (2). We illustrate this through a concrete example.

The single-price representation of the two-asset system. Let n and n' denote two substitutable assets. For ease of exposition, consider the case with homogeneous elasticities across investors – a general case with heterogeneous elasticities is discussed in Chaudhary et al. (2024). Define the flow and price vectors as $\Delta\mathbf{q}_{i,t} \equiv (\Delta q_{i,t}(n),\ \Delta q_{i,t}(n'))^\intercal$, $\Delta\mathbf{p}_t \equiv (\Delta p_t(n),\ \Delta p_t(n'))^\intercal$, and $\mathbf{u}_{i,t} \equiv (u_{i,t}(n),\ u_{i,t}(n'))^\intercal$.

The log-linear demand of investor i is

$$\Delta \mathbf{q}_{i,t} \;=\; \Gamma\,\Delta \mathbf{p}_t + \mathbf{u}_{i,t}, \qquad \Gamma = \begin{pmatrix} -\zeta(n) & \zeta(n,n') \\ \\ \zeta(n',n) & -\zeta(n') \end{pmatrix}, \tag{17}$$

where $\zeta(n)$ is the own-price elasticity for asset n, and $\zeta(n,n')$ is the cross-elasticity of demand for asset n with respect to the price of asset n'.

Using subscript S to denote the size-weighted aggregation, e.g., $\Delta q_{S,t}(\cdot) \equiv \sum_i S_i(\cdot)\,\Delta q_{i,t}(\cdot)$, market clearing gives:

$$\mathbf{0} \;=\; \Gamma\,\Delta \mathbf{p}_t + \mathbf{u}_{S,t}, \qquad \Rightarrow \qquad \Delta \mathbf{p}_t = -\,\Gamma^{-1}\,\mathbf{u}_{S,t}, \tag{18}$$

Inverting the elasticity matrix Γ, the own-price impact of a demand shift to asset n is:

$$\mathcal{M}(n) \;=\; \frac{\partial\,\Delta p_t(n)}{\partial u_{S,t}(n)} \;=\; \frac{1}{\zeta(n)\,(1-\mathcal{Q}_{n\leftarrow n'}\,\mathcal{Q}_{n'\leftarrow n})} \tag{19}$$

where $\mathcal{Q}_{n\leftarrow n'}$ and $\mathcal{Q}_{n'\leftarrow n}$ are defined as

$$\mathcal{Q}_{n\leftarrow n'} := \frac{\zeta(n,n')}{\zeta(n')}, \qquad \mathcal{Q}_{n'\leftarrow n} := \frac{\zeta(n',n)}{\zeta(n)}. \tag{20}$$

We term these coefficients *demand pass-throughs*: $\mathcal{Q}_{n\leftarrow n'}$ measures how a unit demand shift for asset n' translates into an effective demand shift for asset n through substitution.[10]

From the second row of the market-clearing condition (18), we can express the price change of asset n' as a function of the price change of asset n and the demand shift to asset n':

$$\Delta p_t(n') \;=\; \frac{\zeta(n',n)}{\zeta(n')}\,\Delta p_t(n) \;+\; \frac{1}{\zeta(n')}\,u_{S,t}(n'). \tag{21}$$

Substituting (21) into the demand equation for asset n yields the single-price representation:

$$\Delta q_{i,t}(n) \;=\; -\underbrace{\zeta(n)\;(1-\mathcal{Q}_{n\leftarrow n'}\,\mathcal{Q}_{n'\leftarrow n})}_{\tilde{\zeta}(n)\,=\,1/\mathcal{M}(n)}\,\Delta p_t(n) \;+\; \underbrace{\mathcal{Q}_{n\leftarrow n'}\,u_{S,t}(n') \;+\; u_{i,t}(n)}_{\tilde{u}_{i,t}(n)}, \tag{22}$$

[10]Demand pass-throughs relate to cross-elasticities but are expressed in quantity units. While $\zeta(n,n')$ measures how a *price* change in n' affects demand for n, dividing by $\zeta(n')$ converts this to how a *demand shift* in n' affects demand for n. One crucial difference is that cross-elasticities are a pure partial-equilibrium concept, while demand pass-through depends on how the substitute price responds to the demand shift, and hence is intrinsically a general-equilibrium object.

Since (22) mirrors the single-asset demand curve (2), Theorem 1 applies directly with identical construction. The differences lie in interpreting the bound's two key components.

First, while Theorem 1 continues to bound the own-price impact $\mathcal{M}(n)$, the price impact no longer equals the reciprocal of the own-price elasticity $\zeta(n)$; instead, it includes the amplification factor $(1 - \mathcal{Q}_{n \leftarrow n'} \mathcal{Q}_{n' \leftarrow n})$. The distinction arises because price elasticity $\zeta(n)$ measures partial-equilibrium responses – how demand responds to price changes *with substitute prices held fixed.* In contrast, price impact $\mathcal{M}(n)$ captures the feedback loop from the general-equilibrium effects: a demand shift for asset n moves not only its own price but also substitute prices, which recursively feed back into demand for asset n itself.

Second, investor agreement ρ now encompasses both agreement on asset-specific demand shifts $u_{i,t}(n)$ and agreement on substitution effects from the aggregate demand shift to substitutes $u_{S,t}(n')$. The latter enters asset n's demand because it moves substitute prices, effectively acting as a common demand shift for asset n through cross-asset substitution.

When substitution effects remain moderate (small pass-throughs $\mathcal{Q}_{\cdot,\cdot}$), the interpretive differences between single-asset and multi-asset settings are minor. However, strong substitution patterns require careful interpretation of the bound. Appendix C provides detailed analysis and numerical examples for such cases.

3 Data and Empirical Facts

3.1 Data Sources and Variable Construction

Data. Our empirical analyses are all at the quarterly frequency. We obtain quarterly institution-level share holdings $Q_{i,t}(n)$ from 1990 to 2024 from the Thomson Institutional Holdings Database (s34 file). Institutions are denoted by $i = 1, ..., I$. The subscript t indicates the report date of the 13F filing. [11] Further details can be found in Appendix D.1. Subsequently, we merge quarterly stock holdings with data on prices and fundamentals from CRSP, Compustat, and IBES. We restrict our sample to common ordinary shares (share codes 10 and 11) traded on the NYSE, AMEX, and NASDAQ (exchange codes 1, 2, and 3), that have (on average) at least 10 institutional holders and at least 30%

[11] In the main text, we use holdings at the institution level (e.g., BlackRock as a single entity rather than as individual funds) to achieve the most comprehensive coverage. Since holdings are aggregated across funds within the same asset manager, transactions among funds within the same institution are not observed at this level, which could potentially lead to an underestimation of portfolio turnover. However, in Appendix D.1, we show that portfolio turnover computed at the mutual fund level is very close to that at the institution level, suggesting that within-fund-family trades are negligible.

observed institutional ownership.[12] Δ denotes quarterly changes. Ownership shares (size-weights) are denoted by $S_{i,t}(n) = \frac{Q_{i,t}(n)}{\bar{Q}_t(n)}$, where $\bar{Q}_t(n)$ are the total shares outstanding of the stock. Empirically, we do not observe the holdings of *all* investors, but are restricted by reported 13F filings. We therefore construct the trades of the residual sector that holds the remaining shares outstanding such that the trades of all investors sum to 0.[13]

Estimating volatilities. As discussed in Section 2, our bound holds both in sample as well as period by period. We estimate both $\sigma_q(n)$ and $\sigma_p(n)$ in the time series for each stock using five-year backward-looking rolling windows, preventing our results to suffer from forward-looking bias. We estimate $\sigma_p(n)$ using the time-series volatility of quarterly stock returns. As described in Section 2.5, $\sigma_q(n)$ can either be measured directly as $\sqrt{\sum_{i=1}^{I} S_{i,t}\widehat{\text{Var}}(\Delta q_{i,t}(n))}$, the size-weighted average of investor-specific volatilities (the $\mathcal{L}_2$ norm), or approximated using portfolio turnover $\sqrt{\frac{\pi}{2}}\hat{\mathbb{E}}\left[\text{Turnover}_t(n)\right]$ (the $\mathcal{L}_1$ norm). We construct both measures and find similar results. Generally, we favor portfolio turnover for several reasons. First, it is straightforwardly constructed and closely linked to gross trading volume. Second and more importantly, $\mathcal{L}_2$ norms, such as the standard deviation, are susceptible to outliers – a common feature in flow data – while $\mathcal{L}_1$ norms, such as the mean absolute deviation, are more robust estimators of statistical dispersion in the presence of fat tails (due to frequent extensive margin trades).

Unlike σ_q and σ_p, which are directly observable from trade and price data, investor agreement ρ is inherently unobserved. To that end, we first present results that are agnostic about the level of investor agreement. Later in Section 4.2, we present and discuss different strategies of how to empirically measure ρ.

The top panel of Table 1 reports $\sigma_p(n)$ and $\sigma_q(n)$ (both measured via $\mathcal{L}_1$ and $\mathcal{L}_2$ norms). The average share in our sample has a quarterly return volatility $\sigma_p(n)$ of 22%. The average σ_q constructed from portfolio turnover is 25%. The 5th percentile, median, and 95th percentile are given by 8%, 23%, and 50%, respectively. In contrast, the $\mathcal{L}_2$ measure of σ_q is distributed very similarly with a slightly higher average of 30% and the 5th percentile, median, and 95th percentile given by 10%, 30%, and 53%, respectively. As a consequence, the ratio of return volatility to portfolio turnover (hereafter, p/q volatility ratio) equals 1.15 for the average share. However, there exists considerable variation in this ratio as can be seen from the 5th and 95th percentiles, which equal 0.36 and 3.02, respectively.

[12] All results are robust to alternative cut-offs.

[13] All results in the paper are robust to omitting the residual sector and constructing $\bar{Q}_t(n)$ (and the corresponding size weights) as the sum of institutional shares held. However, we prefer to construct the residual sector as doing so effectively accounts for trades by the institutional sector as a whole, which would be omitted otherwise.

Finally, Table 1 also reports moments of the distributions of institutional ownership and trading volume. For example, the average share is held by about 200 institutions with an average institutional ownership share of 60%. Notably, all our main results are robust to restricting the sample to stocks for which institutional ownership is higher than 90%.

Table 1: **Summary Statistics**
The table summarizes the distribution of the key variable inputs for deriving the price impact bounds over the cross-section of US equities. The first rows report the volatility of returns σ_p and the volatility of flows σ_q, both explicitly computed via size-weighted investor-specific volatilities, and the $\mathcal{L}_1$ approximation from scaled portfolio turnover. The volatilities are computed over 5-year rolling windows. The middle panel reports the number of investors holding each stock, the distribution of institutional ownership and the investor concentration defined as $\sum_i S_{i,t}^2(n)$. The last rows report gross quarterly trading volume (from CRSP) alongside portfolio turnover divided by 2. The division by 2 avoids double-counting trades and ensures comparability to gross trading volume.

	Mean	Std	5th pctl.	Median	95 pctl.
Volatilities of Trade and Returns					
Return Volatility σ_p	0.22	0.15	0.09	0.19	0.47
Flow Volatility σ_q ($\mathcal{L}_2$)	0.30	0.13	0.10	0.30	0.53
Portfolio Turnover σ_q ($\mathcal{L}_1$)	0.25	0.13	0.08	0.23	0.50
p/q Volatility Ratio σ_p/σ_q	1.16	1.32	0.36	0.81	3.02
Ownership Distribution					
Number of Institutional Holders	202.94	276.54	14.00	124.00	663.00
Institutional Ownership	0.60	0.25	0.16	0.62	0.99
Ownership Concentration (HHI)	0.23	0.21	0.04	0.16	0.67
Trading Volume					
CRSP Total Volume	0.46	0.55	0.06	0.31	1.35
$\frac{1}{2}$ Net Volume	0.10	0.09	0.02	0.08	0.24

3.2 Portfolio Turnover versus Gross Trading Volume

A well-known feature of equity markets is the high volume of trading. In fact, in the past 30 years, the quarterly trading volume (relative to shares outstanding) for the median stock on the NYSE, AMEX, and NASDAQ, has been around 50-100%. Instead, portfolio turnover constructed from changes in institutional investors' portfolios is considerably lower. For example, as of 2024, the quarterly portfolio turnover amounts to 8% of shares outstanding for the average stock.

Institutional ownership in the average stock in our sample is 60%. We first confirm that the large gap between gross trading volume and (institutional) portfolio turnover is not simply driven by offsetting trades within the unobserved residual investor. Panel a) of Figure 1 shows the ratio of gross trading volume to portfolio turnover across stocks with varying levels of institutional ownership. The ratio remains essentially unchanged across ownership groups, indicating that the gap between gross volume and turnover is not explained by unobserved portfolio adjustments. Even in the top decile – where average institutional ownership reaches 99.7% – gross quarterly trading volume is still about seven times larger than portfolio turnover.

Studies using household data further confirm that households have even smaller portfolio turnover compared to institutional investors. Using portfolio holdings data from households, Gabaix et al. (2025) measure the risk transfer – defined as the percent change in market risk exposure for a group of investors over a given period, a measure closely related to portfolio turnover at the aggregate market level. They find that the quarterly risk transfer is only 0.65% for household groups, far smaller than the 6% portfolio turnover observed for institutional investors at the aggregate market level (as reported in Figure 7 below).

We also verify that the low portfolio turnover is not driven by aggregation across mutual funds within a management company. In Appendix D.1, we disaggregate 13F managers into their constituent mutual funds and ETFs and show that portfolio turnover computed at the fund level is only marginally larger than at the institutional level, suggesting that within-fund-family trading is negligible.

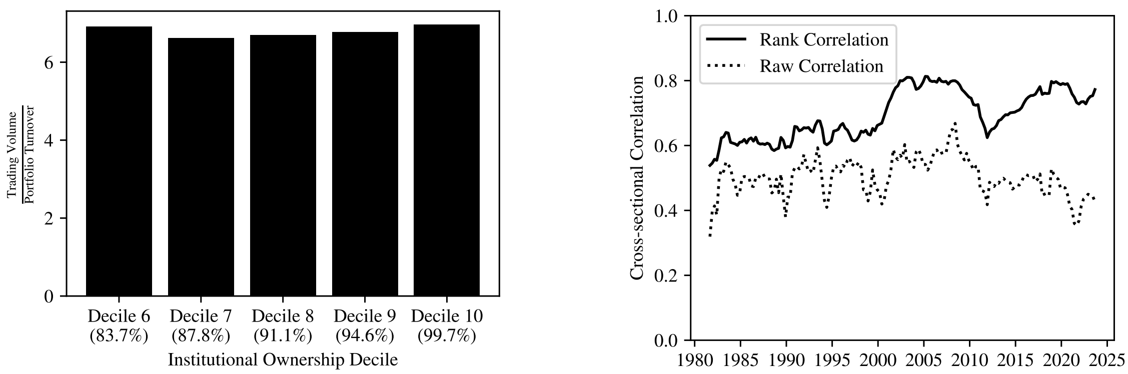
Given the large differences between gross trading volume and portfolio turnover, one may wonder whether these measures are fundamentally economically different. Panel b) of Figure 1 suggests they are not. Despite a difference in levels up to a factor of seven, gross trading volume and portfolio turnover are highly correlated in the cross-section. The cross-sectional correlation in ranks is about 80% in recent periods.[14] This high correlation suggests that portfolio turnover and gross trading volume are at least to some extent driven by the same fundamental economic primitives. The key difference is that gross trading volume is substantially inflated by round-trip trades that do not contribute to long-term liquidity provision.

[14]Mechanically, portfolio turnover is part of gross trading volume by construction. However, this cannot explain the high correlation due to the large discrepancy in the level.

Figure 1: **Portfolio Turnover versus Gross Trading Volume**
Panel a) of the figure plots the ratio of quarterly trading volume (CRSP) relative to portfolio turnover for groups of stocks sorted by institutional ownership. Average institutional ownerships for each decile are reported in brackets. Panel b) plots the quarterly (rank) correlation of trading volume and portfolio turnover in the cross-section of stocks. We report annual averages of quarterly correlations.

(a) $\frac{\text{Trading Volume}}{\text{Portfolio Turnover}}$ by Institutional Ownership

(b) Correlation: Volume vs. Portfolio Turnover



If institutional investors were the only investors trading the underlying securities, then any difference between gross trading volume and portfolio turnover is due to offsetting round-trip trades within a quarter. For example, a high-frequency market maker may hold a small inventory from beginning to end of quarter, but account for a large share of total trading volume in a given stock. Importantly, such high-frequency round-trip trades cannot accommodate persistent demand shifts in the long-term. According to our theory, what matters for the long-term impact of persistent demand shifts is portfolio turnover, not gross trading volume. Dissecting the difference between gross trading volume and institutional turnover and analyzing the effects of high-frequency intermediation for long-term asset pricing is beyond the scope of this paper but an exciting avenue for future research.[15]

4 The Price Impact Bound

4.1 The Price Impact Bound under Varying Levels of Investor Agreement

As discussed above, in a first step, we evaluate the price impact bounds without taking a stance on the level of investor agreement ρ. In particular, we document the bounds $\mathcal{M}(\rho)$ as a function of ρ for U.S. equities. That is, using stock-level return volatility, σ_p, and portfolio turnover, σ_q, constructed as described in the previous section, we compute the bound $\mathcal{M}(\rho)$ for each stock while allowing ρ to vary

[15] In follow-up work, we link the long-term asset pricing implications of portfolio turnover and the short-term microstructural implications of high-frequency trading volume. We show that the market participants that have entered since 1980 – such as high-frequency market makers, ETF authorized participants, algorithmic trading firms, and (mobile) retail traders – have contributed to the surge in trading volumes and a decline of short-term price impact, but did not help absorbing long-term demand shifts.

between 0 and 1. Panel a) of Figure 2 plots the distribution of the lower bounds of price impact for individual U.S. stocks. In contrast, Panel b) plots the distribution of the upper bounds on aggregate elasticity, i.e., the inverse of the price impact bounds.

Figure 2: **Price Impact Bounds under Varying Agreement ρ**
The figure plots the price impact bound for a given level of investor agreement ρ. Panel a) plots the lower bound on the price impact $\mathcal{M}(\rho)$ as a function of ρ, for the average US stock, as well as the top and bottom 10% of stocks with the highest and lowest p/q volatility ratio $\frac{\sigma_p}{\sigma_q}$. The lower bound on price impact bound can be inverted to obtain an upper bound on the aggregate (size-weighted) elasticity. Panel b) plots the upper bound on the aggregate elasticity $\zeta_S(\rho)$ for US stocks.

(a) Price Impact Bound $\mathcal{M}(\rho)$ (b) Elasticity Bound $\zeta_S(\rho)$

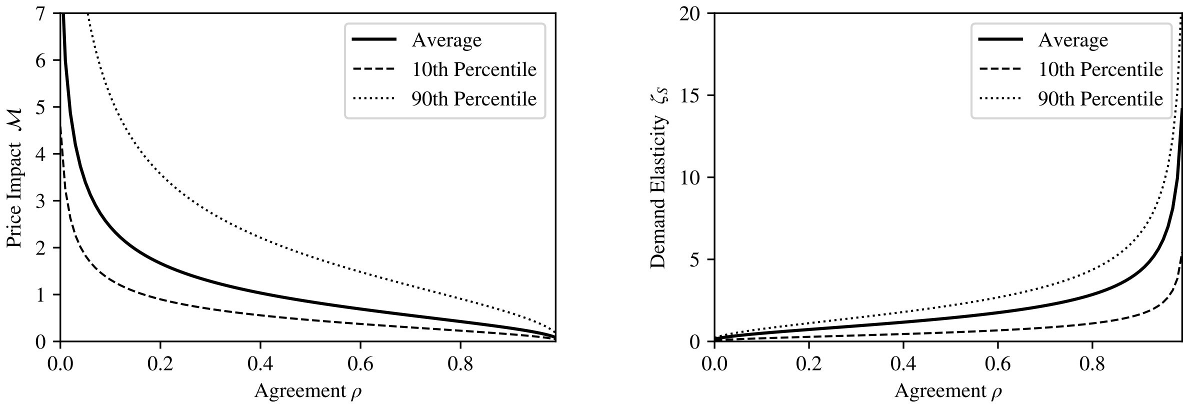


Figure 2 shows that a high p/q volatility ratio is consistent with perfectly elastic markets that feature a close-to-zero price impact. However, such coexistence requires a high degree of agreement among investors, implying that their demand shifts are almost perfectly correlated. Note that the average level of the p/q volatility ratio $\frac{\sigma_p}{\sigma_q}$ is about 1. That is, for a 1% demand shift to move prices less than 0.1% ($\mathcal{M} < 0.1$), investor agreement must exceed 99%. In other words, the empirical level of the p/q volatility ratio can only be reconciled with elastic markets if investors are homogeneous to an unrealistic degree. Put differently, under reasonable levels of investor disagreement, price impact will exceed 0.1%. Generally, the bounds are more stringent as investor agreement decreases ($\rho \to 0$).

4.2 Measuring Investor Agreement

Without an explicit measure of investor agreement, ρ, the lower bound of price impact cannot be determined. However, as can be seen from Figure 2 the bound is relatively flat when ρ lies between 0.2 and 0.8. In this region, the bound varies predominantly due to variation in the p/q volatility ratio – at least in our canonical application to the cross-section of U.S. stocks. Panel a) of Appendix Figure E.2 reinforces this conclusion more formally by plotting the partial derivative of the bound with respect to

ρ. The derivative is small in absolute terms in a large surrounding neighborhood of $\rho = 0.5$ but grows significantly for extreme degrees of agreement/disagreement, i.e., as ρ approaches 1 or 0. The fact that the partial derivative, d, is mostly small implies that a very precise estimate of ρ is required to differentiate between models with $\mathcal{M} = 1$ versus models with $\mathcal{M} = 0.5$. On the other hand, rejecting the null hypothesis that $\mathcal{M} < 0.1$, as implied by most canonical frictionless models, merely requires showing that $\rho < 0.99$. Arguably, this is a much lower hurdle to cross given the extensive literature on heterogeneity in preferences and beliefs among investors. Therefore, rather than trying to provide a precise estimate of ρ, our objective is to establish that ρ is unlikely to be close to either 0 or 1. In this case, the part of Equation (11) which relates to investor heterogeneity, $\mathcal{D} = \sqrt{\frac{1}{\rho} - 1}$, is relatively close to 1 and, thus, the simple p/q volatility ratio is a close proxy for the actual price impact bound.

Unfortunately, common portfolio-based measures of disagreement, such as short interest and active share, cannot directly inform us about investor agreement, as these measures are endogenous to prices and thus already contain information about elasticities – the very quantity we seek to measure. For this reason, we estimate ρ directly from survey data. To that end, we estimate investor agreement via equity analyst agreement. This approach is almost model-free, as it does not require imposing any specific covariance structure on the underlying demand shifts. However, it does require that the estimated ρ for analysts is "portable" and, thus, reflects well the ρ of investors. Importantly, we do not assume that investors and analysts are the same agents – only that the cross-sectional dispersion in analyst forecasts is a reasonable proxy for the heterogeneity in investors' demand shifts. Notably, because analysts tend to operate within a relatively homogeneous professional environment, and because belief disagreement captures only one aspect of broader investor heterogeneity, the limited dispersion in analyst expectations likely overstates the degree of correlation in demand shifts among the full set of investors.

Since analysts submit forecasts across different horizons – from one-quarter ahead to long-term growth rates – and investors care about the total discounted cash flows when trading stocks, we estimate analyst agreement at different horizons. To be consistent with our theoretical framework, we focus on agreement in quarterly forecast updates from Institutional Broker Estimates System (I/B/E/S) stock analysts. Specifically, let $\Delta f^h_{i,t}(n)$ denote the update made by analyst i in period t to the earnings per share (EPS) forecast of firm n for horizon h. We then estimate analyst agreement $\rho^h_{EPS}(n)$ for each stock n and forecast horizon h as the adjusted R^2 from regressing $\Delta f^h_{i,t}(n)$ on time fixed effects:[16]

[16]We first demean forecast updates across time within each analyst, ensuring that the total variation in the regression excludes heterogeneity in average forecast updates. See Appendix D.3 for more details.

$$\Delta f_{i,t}^h(n) = \gamma_t + \epsilon_{i,t}^h(n) \quad \text{for each } n \text{ and } h, \tag{23}$$

where γ_t denotes time fixed effects. We estimate Equation (23) for horizons ranging from one-quarter ahead to three-quarters ahead, as well as long-term growth rates (LTG). The details of the sample construction and estimation procedures can be found in Appendix D.3.

Table 2: Summary Statistics of ρ Estimated from Earnings Forecast Updates
The table reports the distribution of investor agreement ρ estimated from analyst forecast updates using Equation (23). For each stock and forecast horizon, ρ is computed as the adjusted R^2 from regressing analyst forecast updates on time fixed effects. 1Q, 2Q, and 3Q refer to one-quarter ahead, two-quarter ahead, and three-quarter ahead earnings per share (EPS) forecasts, respectively. LTG refers to long-term growth forecasts.

Horizon	**# Firms**	**Mean**	**5th Pctl**	**Median**	**95th Pctl**
1Q	754	0.53	0.16	0.56	0.80
2Q	669	0.47	0.11	0.48	0.76
3Q	585	0.41	0.07	0.4	0.75
LTG	366	0.29	0.0	0.26	0.72

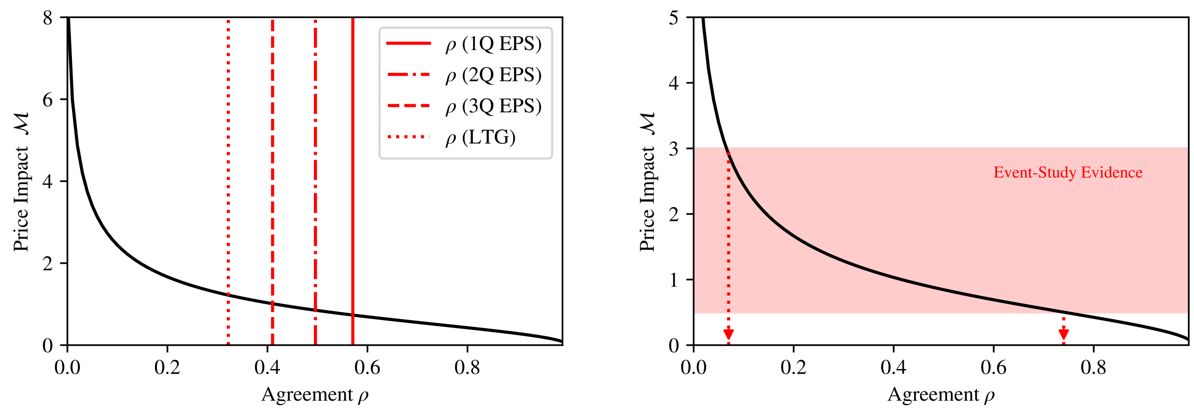
Table 2 reports the cross-sectional distribution of $\rho^h(n)$ for different forecast horizons. Intriguingly, analyst agreement exhibits a clear term structure across forecast horizons: as the horizon increases, analysts increasingly disagree with each other. This pattern is intuitive – forecast uncertainty grows with the forecasting horizon, and fewer reliable common signals are available for analysts to anchor their expectations. Since a stock's value reflects discounted cash flows across all horizons, investors' demand shifts incorporate innovations to expected cash flows possibly across all horizons. Consequently, estimates derived from forecasts for one quarter and long-term growth can be interpreted as lower and upper bounds of investor agreement originating from cash flow expectations.

For stocks in the United States, the average update in one-quarter ahead earnings per share forecasts across analysts explains approximately 53% of the total variation in EPS updates. At this level of investor agreement ($\rho = 53\%$), we obtain an average stock-level price impact of 0.75. In contrast, the average update in long-term growth forecasts explains only 29% of the total variation in LTG updates, implying a price impact of 1.0. Across all measures of investor agreement, we rarely observe values of ρ exceeding 80%, suggesting that for the vast majority of stocks, price impact exceeds 0.5 and price elasticity is below 2.

Event-Study Implied Agreement. Alternatively, we can compare our estimates of investor agreement from the I/B/E/S data against the investor agreement *implied* by event-study estimates of price impact. To that end, we take the empirical estimates of $\mathcal{M}$ at face value and use Equation (11) to impute ρ. Panel a) of Figure 3 reports average ρ based on survey data along with the stock-level price impact bound as a function of ρ for the average stock. We plot the average ρ obtained from 1, 2, and 3-quarter ahead EPS forecast updates, as well as LTG updates. Panel b) documents the implied ρ based on price impact estimates from the literature. For the range of price impacts found in event studies (such as stock index inclusions, mutual fund flow-induced trades, and dividend reinvestments) our bound implies that investor agreement ρ should roughly lie between 0.1 and 0.75. Notably, all our estimates from the I/B/E/S data are well within this range. Last, in Appendix D.4 we examine the investor agreement implied by a structural asset pricing model designed to match both prices and investor-level holdings data. To this end, we use the model by Koijen and Yogo (2019) and estimate the stock-level agreement implied by the demand curves within their framework. We again confirm that $\rho(n)$ is not pathologically high. The average ρ implied by logit demand lies between 0.22 and 0.37.[17]

Figure 3: **Investor Disagreement: IBES versus Event Studies**
Panel a) plots the average ρ from survey data along with the stock-level price impact bound as a function of ρ for the average stock. We plot the average ρ obtained from 1, 2, and 3-quarter ahead EPS forecast updates, as well as LTG updates. Panel b) plots the investor agreement implied from the range of price impacts found in event studies. The dotted lines indicate the implied investor agreement by the event-study range.

(a) Average ρ from Survey Data (b) Implied ρ by Event-Study Literature



4.3 The Price Impact Bound

Next, we apply our estimates of investor agreement to obtain a lower bound on the price impact for each individual stock as follows:

[17] The ρ obtained from logit demand can only inform our bounds to a limited extent, as it requires assuming that logit elasticities from portfolio holdings in levels capture quarterly flow elasticities, which may be violated if investors are inert van der Beck (2022).

$$\mathcal{M}_{\text{EPS}}(n) \equiv \frac{\sigma_p(n)}{\sigma_q(n)}\sqrt{\frac{1}{\rho_{\text{EPS}}(n)} - 1} \tag{24}$$

All of the following results are robust to using any of the four forecast-horizon specific estimates of ρ derived from the I/B/E/S data in our calculations of the lower bound. However, to maximize the cross-sectional sample size, we rely on ρ estimated based on one-quarter ahead EPS forecasts in our baseline results.

As discussed earlier, the EPS-based price impact bound is imperfect, as it ignores investor disagreement along many other dimensions. Moreover, for values of ρ in the neighborhood of 0.5, the term $(\sqrt{\frac{1}{\rho} - 1})$ is close to 1 in magnitude and relatively insensitive to changes in ρ. Therefore, we also consider a simplified bound $\tilde{\mathcal{M}}(n)$ defined as the p/q volatility ratio, implicitly assuming that $\rho(n) = 0.5$ for all stocks n. Formally,

$$\tilde{\mathcal{M}}(n) \equiv \frac{\sigma_p(n)}{\sigma_q(n)}. \tag{25}$$

Henceforth, we refer to $\tilde{\mathcal{M}}(n)$ as the p/q volatility ratio, or simply as the "volatility ratio" when a shorter expression is more convenient. In all our empirical tests, we report results for both $\mathcal{M}_{\text{EPS}}(n)$ and $\tilde{\mathcal{M}}(n)$. Interestingly, $\mathcal{M}_{\text{EPS}}(n)$ contains important incremental information compared to $\tilde{\mathcal{M}}(n)$ when explaining price reactions, despite ρ being measured with noise. Importantly, many other liquidity measures that depend on prices and trading volume, such as Amihud (2002) (and the large body of work that builds on it), are in theory similarly affected by investor agreement, ρ. However, this does not directly become evident as many empirical measures of liquidity do not have a micro-founded equilibrium interpretation.

Panel a) of Figure 4 plots the distribution of the price impact bound $\mathcal{M}_{\text{EPS}}(n)$. For the average stock, the lower bound on the price impact is around 1. The top 5% of stocks have bounds exceeding 3.[18] Overall, there is considerable heterogeneity in the bound across stocks which we will explore in the next section. Importantly, the magnitudes of our bounds are consistent with empirical reduced-form evidence from index inclusions (e.g., Shleifer (1986)), mutual fund flow-induced trades (e.g., Lou (2012)), benchmarking intensity (e.g., Pavlova and Sikorskaya (2022)), and dividend reinvestments (e.g., Schmickler (2020)). Our bounds highlight that low demand elasticities are not an artifact of unique event studies but are instead a pervasive fact that can be directly inferred from the high p/q volatility ratio and the amount of investor disagreement in the market.

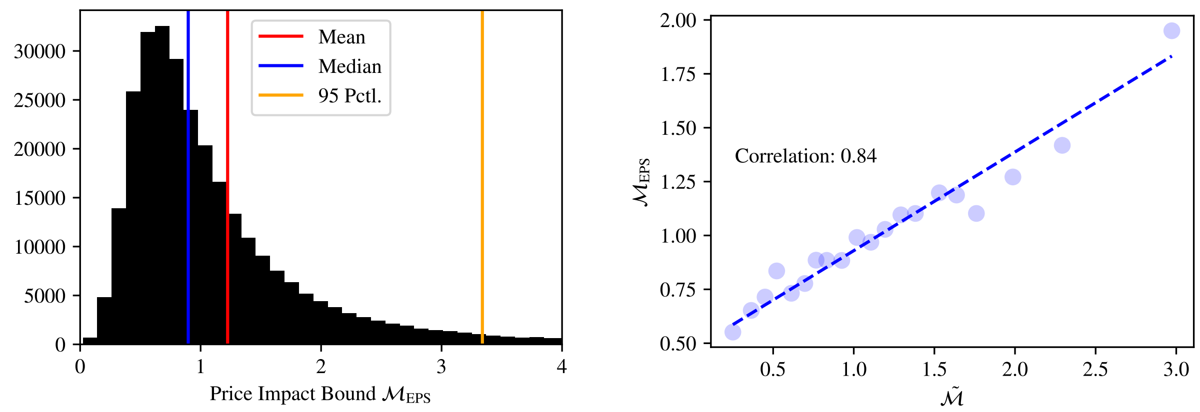
Finally, Panel b) of Figure 4 graphically illustrates the very high cross-sectional correlation between

[18]The distribution of the simple p/q volatility ratio, $\tilde{\mathcal{M}}(n)$, is very similar in shape and magnitudes.

$\tilde{\mathcal{M}}(n)$ and $\mathcal{M}_{\text{EPS}}(n)$ of 84%. Relatedly, Panel b) Appendix Figure E.2 decomposes the cross-sectional variation in $\mathcal{M}_{\text{EPS}}(n)$ and shows that ρ plays a minor role relative to σ_q and σ_p. Put differently, the cross-sectional variation of $\rho_{\text{EPS}}(n)$ is not large enough to dominate the cross-sectional variation in $\frac{\sigma_p(n)}{\sigma_q(n)}$.[19]

Figure 4: **Implied Price Impact for US Equities**
The figure plots the distribution of price impact for the cross-section of US stocks. Panel a) plots the distribution of $\mathcal{M}_{\text{EPS}}(n) \equiv \frac{\sigma_p(n)}{\sigma_q(n)}\sqrt{\frac{1}{\rho_{\text{EPS}}(n)} - 1}$. We use our baseline measure of investor agreement ρ extracted from EPS forecast updates ρ_{EPS}. Panel b) plots the correlation between $\mathcal{M}_{\text{EPS}}$ and the simple p/q volatility ratio $\tilde{\mathcal{M}} = \frac{\sigma_p}{\sigma_q}$.

(a) Price Impact Bound $\mathcal{M}_{\text{EPS}}$ (b) Correlation $\mathcal{M}_{\text{EPS}}$ and $\tilde{\mathcal{M}}$



5 Empirical Relevance of the Stock-Level Bounds

Our stock-level bounds are ultimately theoretical constructs that provide lower limits on the expected price impact of investor-specific demand shifts. That is, the bounds are particularly valuable in settings where empirical estimates are unavailable or difficult to obtain. For instance, identifying a source of plausibly exogenous demand shifts to credibly estimate the price impact for broad portfolios, such as the total U.S. equity market, is challenging. Similarly, estimating asset-level price impact is challenging because much of the carefully identified event-study evidence relies on cross-sectional variation and, thus, obtains pooled estimates across assets. However, to trust our model-implied bounds in such a context, it is crucial to verify that the bounds align well with the empirical evidence from settings with credible identification strategies. To that end, we focus on two of the most widely used and verified event studies in empirical asset pricing. Mutual fund flow-induced trades and index inclusions. In particular, we test whether these (plausibly) exogenous demand shifts imply larger price changes for

[19] As discussed in Section 4.2, more formally, the reason for the minor role of investor agreement is that the derivative $\frac{\partial \mathcal{M}}{\partial \rho} = -\frac{1}{2}\tilde{\mathcal{M}}\frac{1}{\rho^2}\sqrt{1/\rho - 1}$ is small as long as ρ does not take extreme values, i.e. 0 or 1.

stocks with a higher price impact bound, $\mathcal{M}$.

5.1 Flow-induced trades

Following Coval and Stafford (2007), Lou (2012), and Edmans et al. (2012), flow-induced trades by mutual funds (FIT) have been a widely used source of (plausibly) exogenous variation in demand. We follow the construction of flow-induced trades by Lou (2012) and relegate details to the Appendix D.5. To test whether stocks with higher price impact bounds have higher FIT returns, we interact FIT with $\mathcal{M}_{\text{EPS}}$. We then run panel regressions of quarterly stock returns onto FIT, the interaction of FIT with our bounds, and time fixed effects. As expected, the impact of flow-induced trades is significantly larger for stocks with higher price impact bounds as evidenced by the positive and statistically significant coefficient on the interaction term. Moreover, we sort the stocks into quintiles based on $\mathcal{M}_{\text{EPS}}$ and estimate the flow-induced price impact for each quintile by interacting FIT with quintile dummies. Panel a) of Figure 5 plots the results graphically and Appendix Table E.1 reports the results numerically. In line with our theoretical predictions, price impact estimates increase monotonically from the lowest to the highest price impact bound quintile. For example, the flow-driven price impact for the top quintile of stocks is about twice as large as in the bottom quintile.

5.2 Index Inclusions

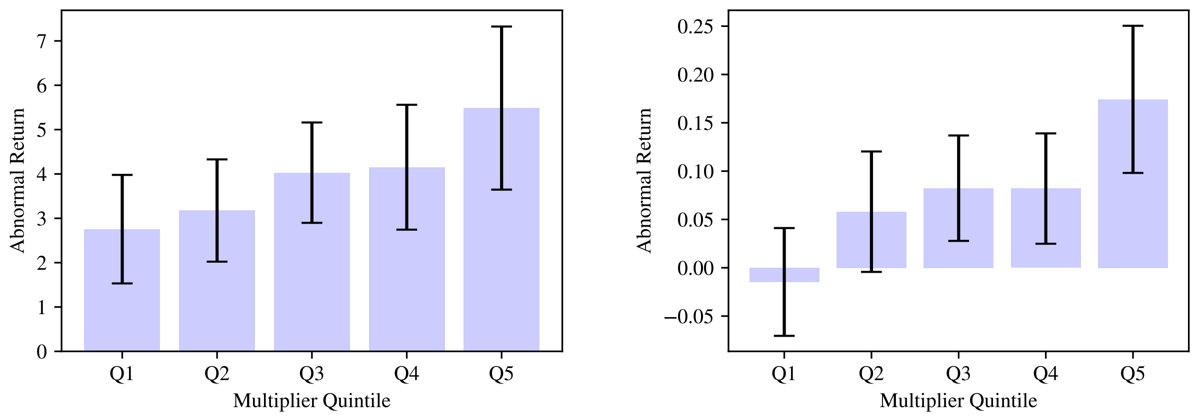
Following Shleifer (1986) and Harris and Gurel (1986), an extensive body of literature investigates the average (abnormal) return around index inclusions and exclusions.[20] Index reconstitutions imply large uninformed demand shifts for affected securities, stemming from passive index trackers who mechanically buy the included and delete the excluded stocks from their portfolios. Relying on the data provided by Greenwood and Sammon (2025) on abnormal event returns and S&P 500 reconstitutions, we find an average abnormal event return of 8%. However, there is considerable variation in event returns with the cross-sectional standard deviation being equal to 12%. Similar to Section 5.1, we examine whether stocks with higher $\mathcal{M}_{\text{EPS}}$ experience significantly higher abnormal event returns. We find that our price impact bounds are highly statistically significantly related to abnormal event returns. In other words, stocks with high bounds have significantly higher (lower) returns when included in (excluded from) the S&P 500. As for flow-induced trades, we sort the included and excluded stocks into quintiles by their price impact bound and regress event returns onto the quintiles. Greenwood and

[20] Among others, Petajisto (2011), Madhavan (2003), Chang et al. (2015), Pavlova and Sikorskaya (2022)

Sammon (2025) find that index returns from announcement to effective reconstitution have declined over time, likely because investors increasingly front-run inclusions ahead of the announcement, spreading the effect over a longer window. We therefore focus on the pre-2000 period, when the average index effect was strongest. We also report the results for the whole sample period, which are quantitatively and qualitatively unchanged, but statistically weaker. Panel b) of Figure 5 plots the results graphically and Appendix Table E.2 reports the results numerically. As before, abnormal returns are increasing when moving from the lowest to the highest price impact bound quintile. For example, the abnormal inclusion return for the top quintile of stocks is about 2.5 times as large as that of the bottom quintile.

Figure 5: **Validation: S&P 500 Inclusions and Flow-Induced Trades**
The figure summarizes the empirical validation of our bounds. Panel a) plots the coefficient of regressing quarterly stock-returns onto flow-induced trades (FIT) interacted with quintile dummies of our price impact bound. Panel b) plots the coefficient of regressing (signed) abnormal event returns during S&P 500 index reconstitutions onto quintile dummies of our price impact bound. We report 95% confidence intervals using standard errors clustered by date.

(a) Mutual Fund Flow Induced Trades (b) S&P 500 Inclusions/Deletions



Lastly, we confirm that the ability of our bounds to price persistent demand shifts is not subsumed by standard high-frequency liquidity measures, which rely on gross trading volume rather than portfolio turnover. In particular, we use two alternative measures: the ratio of return volatility to gross-volume $\frac{\sigma_p}{\text{CRSP Vol.}}$ (as opposed to using portfolio turnover σ_q in the denominator); and the Amihud (2002) illiquidity measure. Appendix Tables E.3 and E.4 repeat the FIT and S&P500 inclusion regressions for these alternative measures. Importantly, the measures based on gross trading volume do *not explain* the abnormal returns due to demand shocks. In fact, relying on gross trading volume rather than portfolio turnover renders the interaction term between FIT and the price impact measure $\mathcal{M}$ insignificant. This suggests that – beyond simply inflating portfolio turnover due to round-trip trades – gross trading volume is less suited to measure *long-term* liquidity provision.

6 Exploring the Bounds Beyond Event Studies

The previous section documented that our bounds are empirically relevant when measuring long-term price impact of investor-specific demand shifts. Based on this evidence, we next explore the cross-sectional variation of our measures in different settings. To do this, we rely on our simplified measure, $\tilde{\mathcal{M}}$, whenever there is no suitable measure of investor agreement ρ available, for example, due to limited time-series variation or lack of estimates of ρ for aggregated portfolios.

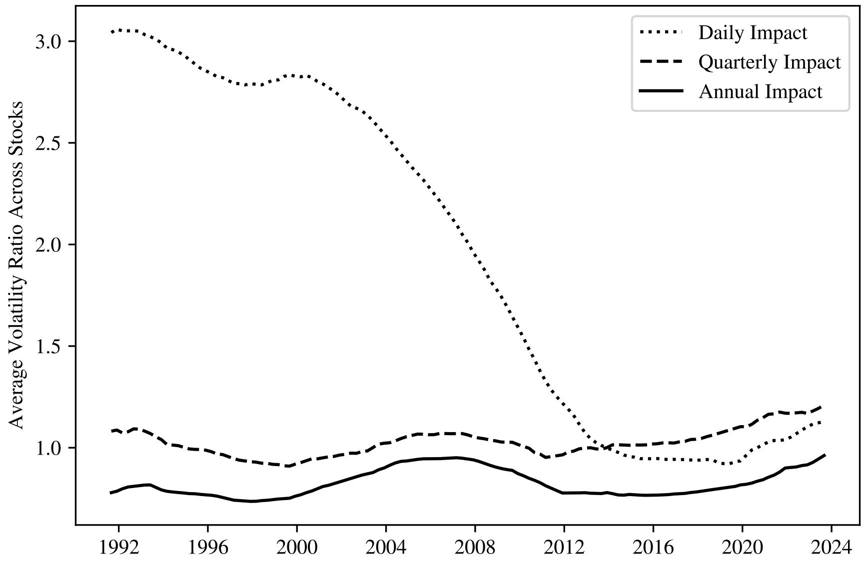
6.1 Price Impact Bound at Different Horizons

A natural feature of financial markets is a decreasing price impact at longer horizons. This pattern is consistent with long-term investors being more willing to bear long-term risks, accepting lower compensation for absorbing risks from higher-frequency traders who initially take the other side of a demand shift (Duffie, 2010). Decreasing price impact, therefore, implies that each layer of intermediation is compensated for providing liquidity. Instead, increasing price impact over time would imply that liquidity providers lose money on average as mispricing amplifies. Our bounds allow testing the extent to which price impacts decrease at lower frequencies and how the relationship between high- and low-frequency impact has evolved over time.

To this end, we construct portfolio turnover and return volatility at different horizons, denoted by $\sigma_{q,H}$ and $\sigma_{p,H}$, where $H = D, Q, Y$ corresponds to daily, quarterly, and yearly measures, respectively. Return volatility at these frequencies is straightforward to compute. Portfolio turnover at lower frequencies (e.g., annual) is also easily constructed by measuring changes in portfolio holdings relative to shares outstanding and summing their absolute values. Estimating portfolio turnover at higher frequencies requires additional assumptions, as comprehensive investor-level holdings data are only available at the quarterly level. However, at an infinitely high frequency, portfolio turnover coincides with gross trading volume. As the frequency decreases, this equivalence breaks down: round-trip trades accumulate, and the gap between portfolio turnover and gross trading volume widens monotonically. To construct the volatility ratio at high (daily) frequency, we make a simplifying assumption and use daily gross trading volume as a proxy for daily portfolio turnover $\sigma_{q,D}$. This effectively abstracts from intra-day round-trip trades, which likely occur because of the role of market makers and high-frequency participants. As a result, daily gross trading volume strictly overstates daily portfolio turnover, implying that the resulting volatility ratio $\sigma_{p,D}/\sigma_{q,D}$ should be interpreted as a conservative (lower-bound) estimate, given that the true $\sigma_{q,D}$ is plausibly lower.

We construct the volatility ratio $\tilde{\mathcal{M}}_H(n) = \frac{\sigma_{p,H}(n)}{\sigma_{q,H}(n)}$ for every stock at the daily, quarterly, and annual level. Figure 6 plots the cross-sectional average for each horizon over time.

Figure 6: **Price Impact at Different Horizons**
The figure plots the stock-level p/q volatility ratio $\tilde{\mathcal{M}}$ for the average stock from 1990 to 2024. $\tilde{\mathcal{M}}$ uses portfolio turnover and return volatility at the daily, quarterly, and annual frequency. We plot ten-year rolling averages of the volatility ratios for visual clarity. The underlying (unaveraged) time series are reported in Figure E.3 in the Appendix.



First, price impact (as proxied by the p/q volatility ratio) decreases monotonically at longer horizons. In 2000, the average daily price impact was over twice as large as the quarterly impact, which in turn was 50% larger than the annual impact.[21] Second, while daily price impact has decreased considerably over our sample period, quarterly and annual price impacts have remained largely flat. This pattern suggests that while markets have become more effective at absorbing demand shocks in the short run, their ability to accommodate long-term shifts has remained largely unchanged. Our focus in this paper lies on the asset pricing implications of persistent, long-horizon quantities, rather than the micro-structural effects of high-frequency trades. Investigating how price impact at different frequencies is connected lies beyond the scope of this paper, but represents an important direction for future research. In particular, it would be interesting to examine whether the rise of high-frequency trading and the resulting improvement in daily liquidity have positive spillover effects on the long-term elasticity faced by investors, for example by lowering their effective trading costs.

We emphasize that inelastic markets and long-term price impact, while economically meaningful, do not imply that price impact increases at longer horizons (which would entail systematic losses for

[21]Note that, because daily price impact is computed using gross volume (rather than portfolio turnover), it represents a lower bound on the true daily volatility ratio.

liquidity providers). Instead, we find that price impact declines monotonically with horizon. This perspective also mitigates concerns that long-term impact is overstated: in fact, the high-frequency estimates reported in Brokmann et al. (2015); Frazzini et al. (2018); Bouchaud et al. (2018) are substantially larger than the bounds we obtain at quarterly and annual frequencies.

6.2 Differences in Price Impact in the Cross-Section of Stocks

In the following, we ask the question: Which stocks have higher price impact bounds? To this end, we regress $\mathcal{M}_{\text{EPS}}$ and $\tilde{\mathcal{M}}$ on various stock-specific characteristics such as size (market equity), systematic risk (market beta), momentum (cumulative past returns), book-to-market ratio, dividend to book equity ratio, profitability, and the illiquidity measure from Amihud (2002). Table 3 reports the results.

Table 3: **Heterogeneity in $\mathcal{M}$**
The table summarizes how $\mathcal{M}$ varies across different stocks. We regress $\mathcal{M}$ and the simplified $\frac{\sigma_p}{\sigma_q}$ on the stock-specific characteristics, log market equity, market beta, momentum, dividend to book equity, profitability, and amihud illiquidity.

	$\mathcal{M}_{\text{EPS}}$			$\tilde{\mathcal{M}}$
	(1)	(2)	(3)	(4)
log(ME)	-0.396***	-0.383***	-0.675***	-0.593***
	(0.017)	(0.017)	(0.033)	(0.029)
β	0.121***	0.159***	0.148***	0.134***
	(0.010)	(0.011)	(0.013)	(0.012)
Cum. Ret.	0.160***	0.151***	0.138***	0.125***
	(0.013)	(0.009)	(0.007)	(0.006)
BM	-0.112***	-0.111***	-0.127***	-0.111***
	(0.013)	(0.012)	(0.012)	(0.011)
$\frac{\text{Dividend}}{\text{BE}}$	0.015	0.023*	-0.028*	-0.026*
	(0.009)	(0.009)	(0.011)	(0.010)
Profit	-0.085***	-0.088***	-0.001	0.001
	(0.011)	(0.010)	(0.011)	(0.010)
Amihud Illiquidity	0.261***	0.265***	0.182***	0.167***
	(0.017)	(0.017)	(0.014)	(0.013)
Date	-	x	x	x
Stock	-	-	x	x
Observations	287895	287895	287895	287895
R^2	0.255	0.283	0.586	0.590
R^2 Within	-	0.250	0.147	0.148

Significance levels: $*$ p $<$ 0.05, $**$ p $<$ 0.01, $***$ p $<$ 0.001. Format of coefficient cell: Coefficient (Std. Error)

First, we find that $\mathcal{M}$ is significantly smaller for larger stocks. That is, a one standard deviation increase in stock size is associated with a 0.13 decline in price impact (t-statistic of 12). This aligns with the view that larger stocks are more liquid, possibly due to more precise and readily available information. Notably, however, this finding contrasts Haddad et al. (2021) and Jiang et al. (2025), who document that large stocks are *less* elastic than small stocks.

Second, stocks with higher market betas exhibit significantly larger price impacts, i.e., a one standard deviation increase in market beta raises the price impact by 0.1. This is consistent with standard CARA-normal intuition: stocks that contribute more to the risk of an arbitrage portfolio are more sensitive to demand shifts (Greenwood (2005), Kozak et al. (2018)).

Third, stocks with stronger past cumulative returns (i.e., momentum stocks) have significantly larger price impacts. This finding aligns with the idea that momentum traders – with upward-sloping demand curves – continue to trade in the direction of the initial price movement, thereby reducing market liquidity and further amplifying price shifts.

Fourth, stocks with higher Amihud (2002) illiquidity have a higher $\mathcal{M}$. Perhaps, this is not surprising as our bounds could be interpreted as low-frequency counterpart to the original Amihud (2002) illiquidity measure. Importantly, however, Amihud (2002) illiquidity does not explain an economically meaningful fraction of our price impact relative to other characteristics. That is, a one standard deviation increase in illiquidity is associated only with an economically relatively small increase in price impact of 0.04. As argued in Section 3.2, gross trading volume (as opposed to portfolio turnover) is not well-suited to assess the price impact of long-term demand shifts.

Importantly, all documented patterns are robust – in fact become stronger – when we additionally control for stock fixed effects. Finally, our results also remain unchanged when we use our simplified bounds, $\tilde{\mathcal{M}} = \frac{\sigma_p}{\sigma_q}$, as an independent variable. This further corroborates the fact that our results appear not to be driven by the investor agreement parameter which is notoriously difficult to quantify.

6.3 Price Impact at Higher Levels of Aggregation

Our bounds are particularly helpful for investigating settings for which there is a lack of relevant and exogenous demand shifts, such as the aggregate stock market. Gabaix and Koijen (2021) find that the aggregate stock market is considerably more inelastic than individual stocks. Li and Lin (2022) find that price multipliers in the cross-section of individual stocks monotonically increase at higher levels of aggregation. Our simplified bounds are informative about the price impact at different levels of aggregation as they rely only on two simple empirical moments: return volatility and portfolio turnover.

Specifically, we compute $\tilde{\mathcal{M}}$ using various aggregation levels. That is, we start from individual stocks and then successively aggregate to 49 Fama-French industry portfolios, 12 Fama-French industry portfolios, the six portfolios double-sorted on size and book-to-market, three portfolios sorted on size, and, finally, one overall market portfolio. To avoid confusion with "portfolio turnover" – which refers specifically to trading activity from institutional holdings – we henceforth refer to these aggregation levels as groups, rather than portfolios. We first compute return volatility $\sigma_p(g)$ and portfolio turnover $\sigma_q(g)$ at these different levels of aggregation. Let $g \subseteq N$ denote the subset of stocks belonging to a given group. Return volatility for group g is then simply the rolling 5-year standard deviation of the value-weighted portfolio return. For example, for the aggregate stock market, $\sigma_p(g)$ is the standard deviation of the value-weighted return across all stocks. Portfolio turnover for group g is given by

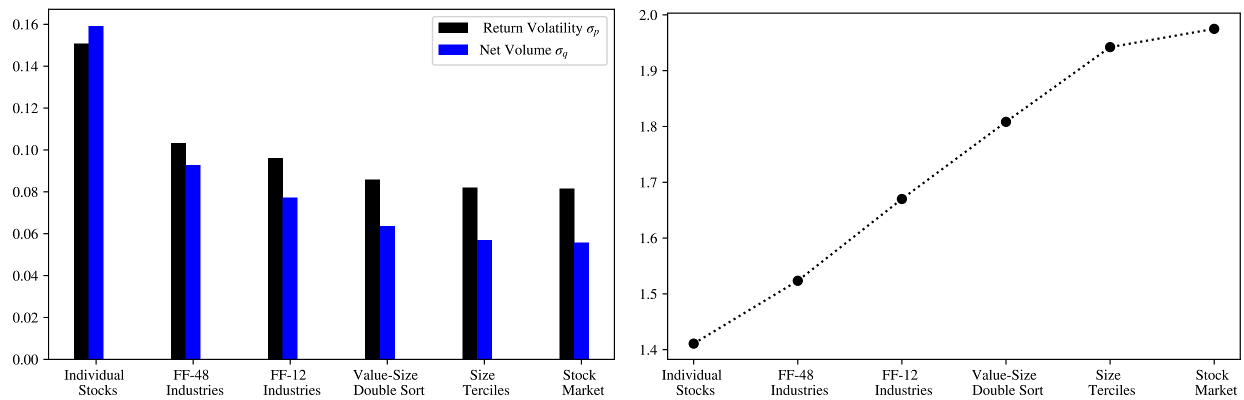
$$\text{Turnover}_t(g) = \frac{\sum_{i=1} \Delta |Q_{i,t}(g)|}{Q_{t-1}(g)}, \tag{26}$$

where $Q_{i,t}(g) = \sum_{n \in g} \Delta Q_{i,t}(n) P_{t-1}(n)$ and $Q_{t-1}(g) = \sum_{i=1}^{I} \sum_{n \in g} Q_{i,t-1}(n) P_{t-1}(n)$. The numerator measures the total dollar flow in and out of group g between $t-1$ and t. The denominator measures the total dollar value of group g as of $t-1$. For example, for the aggregate stock market, the denominator is given by the total stock market capitalization. As before, we then approximate σ_q as the scaled average of portfolio turnover, $\sigma_q(g) \approx \frac{\sqrt{\pi/2}}{T} \mathbb{E}[\text{Turnover}_t(g)]$, estimated from 5-year rolling windows.

Figure 7: **Bounds at different Levels of Aggregation**
Panel a) plots the volatility of returns $\sigma_p(g)$ and portfolio turnover $\sigma_q(g)$ at different levels of aggregation g ranging from individual stocks to the aggregate stock market. Panel b) plots the bound-implied price impact for different levels of aggregation ranging from individual stocks, industries, characteristic portfolios and the aggregate stock market. For each level of aggregation, we plot $\frac{\sigma_p}{\sigma_q}$, which is the price impact implied by $\rho = 0.5$.

(a) Volatility and Turnover

(b) Price Impact



Panel a) of Figure 7 plots our estimates of $\sigma_q(g)$ and $\sigma_p(g)$ at seven different levels of aggregation. At the individual stock level, portfolio turnover σ_q is largest. However, as we aggregate stocks into fewer and fewer buckets, σ_q systematically declines. This pattern is intuitive: Investors' trades in a given stock within an aggregation level partly offset each other, which reduces portfolio turnover. At the same time, return volatility also declines with aggregation. As before, this is intuitive and expected from basic portfolio theory, where diversification reduces idiosyncratic risk. Importantly, however, what matters most for our price impact bounds is the relative speed at which the volatility of returns and portfolio turnover decline – ultimately, an empirical question. In the data, return volatility decreases at a lower pace. As a result, $\tilde{\mathcal{M}}$ rises with aggregation. Panel b of Figure 7 shows that the average $\tilde{\mathcal{M}}$ increases monotonically with the level of aggregation—from 1.4 for individual stocks, to 1.7 for industry portfolios, and up to 2.0 for the aggregate market.

6.4 Price Impact and the Cross-Section of Risk Premia

Although not the main focus of our paper, we also examine whether long-term price impact – as captured by our bounds – predicts expected returns, perhaps because long-term illiquidity constitutes a priced risk. To this end, we conduct simple portfolio sorts. We sort stocks at the beginning of month t into quintiles based on the last available quarterly measure of price impact.[22] Table 4 shows the time series averages of monthly equal-weighted portfolio returns. We report both raw returns and four-factor adjusted returns calculated using the three Fama-French factors (MKT, SMB, and HML) and the momentum factor (UMD). In addition, we also report return spreads between the top and bottom quintiles along with their t-statistics. Both the raw and the four-factor adjusted returns exhibit highly statistically and economically significant return spreads ranging between 0.38% and 0.67% per month across price impact bound measures. Appendix Table E.5 confirms these results in monthly Fama-MacBeth Regressions.

[22] In particular, we transform our quarterly data set to a monthly data set by forward filling the quarterly price impact measures by three months. For example, the relevant measure for the months March 2022, April 2022, May 2022, June 2022, and July 2022 are based on data from the months December 2021, March 2022, March 2022, March 2022, and June 2022, respectively.

Table 4: **Portfolio Sorts**
This table reports monthly returns (in percentage) of portfolios sorted on various price impact bound measures. At the beginning of each month t, stocks are sorted into quintile portfolios according to our measures of month $t-1$. We then calculate monthly equal-weighted portfolio returns for the quintile portfolios and report time-series average portfolio raw returns or four-factor-adjusted returns (MKT, SMB, HML, and UMD), where the factor loadings are estimated in the preceding 60 months. The differences between the top and bottom quintiles are also reported with associated t-statistics. The t-statistics are calculated using Newey-West robust standard errors with six lags.

Sorting Variable	Returns	Low	2	3	4	High	H–L	t-stat
$\tilde{\mathcal{M}}$	Raw Return	1.065	1.099	1.079	1.152	1.688	0.623***	3.746
	Four-Factor-adjusted	0.043	0.058	-0.061	0.077	0.482	0.440***	3.889
$\mathcal{M}_{\text{LTG}}$	Raw Return	0.987	1.034	1.033	1.137	1.646	0.659***	3.756
	Four-Factor-adjusted	0.032	0.066	-0.079	0.090	0.419	0.387***	3.190
$\mathcal{M}_{\text{1Q-EPS}}$	Raw Return	1.003	1.001	1.027	1.146	1.660	0.657***	3.506
	Four-Factor-adjusted	0.025	0.073	-0.090	0.117	0.403	0.378***	2.958
$\mathcal{M}_{\text{2Q-EPS}}$	Raw Return	0.990	1.023	1.036	1.134	1.655	0.665***	3.439
	Four-Factor-adjusted	0.035	0.078	-0.074	0.087	0.402	0.366***	2.877
$\mathcal{M}_{\text{3Q-EPS}}$	Raw Return	1.002	1.020	1.035	1.121	1.659	0.657***	3.460
	Four-Factor-adjusted	0.022	0.083	-0.081	0.072	0.433	0.411***	3.275

Significance levels: $*$ p $<$ 0.1, $**$ p $<$ 0.05, $***$ p $<$ 0.01.

Both the portfolio sorts and Fama–MacBeth regressions show that stocks with higher bounds have significantly higher expected returns in the cross-section. It is important to note, however, that because our theory does not directly speak to expected returns or the pricing of liquidity risk, these results should not be interpreted as an empirical test of the validity of our bounds. Rather, they should be seen as another dimension of their usefulness. This exercise may inform future research on the risks associated with long-term illiquidity as captured by our measure of price impact.

7 Conclusion

This paper reveals a fundamental tension between investor agreement and price impact: when return volatility is high while portfolio turnover is low, market participants cannot simultaneously disagree with each other and respond elastically to price changes. Otherwise we would observe much higher portfolio turnover. This implies that if one acknowledges that investors are not in perfect agreement with each other, one must also concede to considerable price impact, i.e., that markets are inelastic. In other words, given observable moments on quantities and prices, investor agreement and price impact cannot be modeled independently. Highly elastic investors (and thus low price impact) can only be reconciled with the data if investors exhibit a high degree of agreement with one another.

We formalize this trade-off through a model-free bound, $\mathcal{M} \geq \frac{\sigma_p}{\sigma_q} \times \sqrt{\frac{1}{\rho} - 1}$, that connects return volatility, portfolio turnover, and investor agreement to the price impact of persistent demand shifts. Our bounds inform the two competing views on the drivers of asset prices. The first view holds that trading volume is merely a byproduct of price formation and contains no incremental information beyond the representative investor's demand. The second view posits that trading volume is fundamental to understanding price movements, as shifts in quantities interact with a non-zero price impact.

Applied to U.S. equities, our bounds imply substantial price impacts of 0.75 to 1.0 for individual stocks, closely aligning with event study evidence from S&P 500 inclusions and mutual fund flows while traditional high-frequency liquidity measures fail to explain these price impacts. Our bounds vary systematically across assets – with larger stocks exhibiting lower price impacts and higher-beta stocks showing greater impacts – and increase substantially with portfolio aggregation, reaching approximately 2.0 for the aggregate stock market. Our bound provides a diagnostic tool for structural models seeking to reconcile portfolio turnover and return volatilities, and a sanity check for empirical studies on investor disagreement and price impact.

References

Aghaee, A. (2025). The flattening demand curves. *Available at SSRN 4300747*.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets 5*(1), 31–56.

Anshuman, V. R. and S. Viswanathan (2005). Costly collateral and illiquidity. *Duke University working paper*.

Bamber, L. S., O. E. Barron, and T. L. Stober (1999). Differential interpretations and trading volume. *Journal of financial and Quantitative Analysis 34*(3), 369–386.

Banerjee, S. and I. Kremer (2010). Disagreement and learning: Dynamic patterns of trade. *The Journal of Finance 65*(4), 1269–1302.

Barber, B. M. and T. Odean (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The quarterly journal of economics 116*(1), 261–292.

Barber, B. M. and T. Odean (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The review of financial studies 21*(2), 785–818.

Bouchaud, J.-P. (2022). The inelastic market hypothesis: a microstructural interpretation. *Quantitative Finance 22*(10), 1785–1795.

Bouchaud, J.-P., J. Bonart, J. Donier, and M. Gould (2018). *Trades, quotes and prices: financial markets under the microscope.* Cambridge University Press.

Brennan, M., S.-W. Huh, and A. Subrahmanyam (2013). An analysis of the amihud illiquidity premium. *The Review of Asset Pricing Studies 3*(1), 133–176.

Brennan, M. J., T. Chordia, and A. Subrahmanyam (1998). Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics 49*, 345–373.

Brennan, M. J. and A. Subrahmanyam (1996). Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics 41*, 441–464.

Bretscher, L., R. Lewis, and S. Shrihari (2025). Investor betas. *Swiss Finance Institute Research Paper*.

Bretscher, L., L. Schmid, I. Sen, and V. Sharma (2025). Institutional corporate bond pricing. *Review of Financial Studies* (forthcoming).

Bretscher, L., L. Schmid, and T. Ye (2025). Passive demand and active supply: Evidence from maturity-mandated corporate bond funds. *Swiss Finance Institute Research Paper*.

Brokmann, X., E. Serie, J. Kockelkoren, and J.-P. Bouchaud (2015). Slow decay of impact in equity markets. *Market Microstructure and Liquidity 1*(02), 1550007.

Brunnermeier, M. and L. Pedersen (2009). Market liquidity and funding liquidity. *Review of Financial Studies 22*, 2201–2238.

Campbell, J. Y. and L. M. Viceira (2002). *Strategic Asset Allocation: Portfolio Choice for Long-term Investors*. Oxford University Press.

Chang, Y.-C., H. Hong, and I. Liskovich (2015). Regression discontinuity and the price effects of stock market indexing. *The Review of Financial Studies 28*(1), 212–246.

Chaudhary, M., Z. Fu, and J. Li (2023). Corporate bond multipliers: Substitutes matter. *Available at SSRN*.

Chaudhary, M., Z. Fu, and H. Zhou (2024). Anatomy of the treasury market: Who moves yields? *Available at SSRN*.

Chaudhary, M., F. Zhiyu, and J. Li (2024). Corporate bond multipliers: Substitutes matter. *Available at SSRN*.

Chordia, T., A. Subrahmanyam, and V. R. Anshuman (2001). Liquidity shocks and equilibrium liquidity premia. *Journal of Financial Economics 59*, 32.

Constantinides, G. (1986). Capital market equilibrium with transaction costs. *Journal of Political Economy 94*, 842–862.

Coqueret, G. (2022). Characteristics-driven returns in equilibrium. *arXiv preprint arXiv:2203.07865*.

Couts, S. J., A. S Gonçalves, Y. Liu, and J. Loudis (2024). Institutional investors' subjective risk premia: Time variation and disagreement.

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

Dahlquist, M. and M. Ibert (2024). Equity return expectations and portfolios: Evidence from large asset managers. *The Review of Financial Studies 37*(6), 1887–1928.

Darmouni, O., K. Siani, and K. Xiao (2022). Nonbank fragility in credit markets: Evidence from a two-layer asset demand system. *Available at SSRN 4288695*.

Datar, V. T., N. Y. Naik, and R. Radcliffe (1998). Liquidity and asset returns: An alternative test. *Journal of Financial Markets 1*, 203–219.

Davis, S. J. and J. Haltiwanger (1992, August). Gross Job Creation, Gross Job Destruction, and Employment Reallocation*. *The Quarterly Journal of Economics 107*(3), 819–863.

Duffie, D. (2010). Presidential address: Asset price dynamics with slow-moving capital. *The Journal of finance 65*(4), 1237–1267.

Edmans, A., I. Goldstein, and W. Jiang (2012). The real effects of financial markets: The impact of prices on takeovers. *The Journal of Finance 67*(3), 933–971.

Fang, X., B. Hardy, and K. K. Lewis (2022). Who holds sovereign debt and why it matters. Technical report, National Bureau of Economic Research.

Frazzini, A., R. Israel, and T. J. Moskowitz (2018). Trading costs. *Available at SSRN 3229719*.

Gabaix, X. and R. S. Koijen (2021). In search of the origins of financial fluctuations: The inelastic markets hypothesis. *Available at SSRN 3686935*.

Gabaix, X., R. S. J. Koijen, F. Mainardi, S. S. Oh, and M. Yogo (2025). Limited Risk Transfer Between Investors: A New Benchmark for Macro-Finance Models.

Garleanu, N. and L. Pedersen (2007). Liquidity and risk management. *American Economic Review 97*, 193–197.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus (2021). Five facts about beliefs and portfolios. *American Economic Review 111*(5), 1481–1522.

Greenwood, R. (2005). Short-and long-term demand curves for stocks: theory and evidence on the dynamics of arbitrage. *Journal of Financial Economics 75*(3), 607–649.

Greenwood, R. and M. Sammon (2025). The disappearing index effect. *The Journal of Finance 80*(2), 657–698.

Grossman, S. J. and J. E. Stiglitz (1980). On the Impossibility of Informationally Efficient Markets. *The American Economic Review 70*(3), 393–408.

Guiso, L., P. Sapienza, and L. Zingales (2008). Trusting the stock market. *the Journal of Finance 63*(6), 2557–2600.

Haddad, V., Z. He, P. Huebner, P. Kondor, and E. Loualiche (2025, March). Causal Inference for Asset Pricing.

Haddad, V., P. Huebner, and E. Loualiche (2021). How competitive is the stock market? theory, evidence from portfolios, and implications for the rise of passive investing. *Theory, Evidence from Portfolios, and Implications for the Rise of Passive Investing (April 7, 2021)*.

Han, X., N. L. Roussanov, and H. Ruan (2021). Mutual fund risk shifting and risk anomalies. *Available at SSRN 3931449*.

Hansen, L. P. and R. Jagannathan (1991, April). Implications of Security Market Data for Models of Dynamic Economies. *Journal of Political Economy 99*(2), 225–262.

Harris, L. and E. Gurel (1986). Price and volume effects associated with changes in the s&p 500 list: New evidence for the existence of price pressures. *the Journal of Finance 41*(4), 815–829.

Harris, M. and A. Raviv (1993). Differences of opinion make a horse race. *The Review of Financial Studies 6*(3), 473–506.

Hartzmark, S. M. and D. H. Solomon (2021). Predictable price pressure. *Available at SSRN 3853096*.

Heaton, J. and D. J. Lucas (1996). Evaluating the effects of incomplete markets on risk sharing and asset pricing. *Journal of Political Economy 104*, 443–487.

Hellwig, M. F. (1980, June). On the aggregation of information in competitive markets. *Journal of Economic Theory 22*(3), 477–498.

Hong, H. and J. C. Stein (2007). Disagreement and the stock market. *Journal of Economic perspectives 21*(2), 109–128.

Honkanen, P., Y. Zhang, and A. Zhou, T (2025). Etf dividend cycles predict money market fund flows and treasury yield changes. *Critical Finance Review Forthcoming*.

Huang, M. (2003). Liquidity shocks and equilibrium liquidity premia. *Journal of Economic Theory 109*, 104–129.

Huebner, P. (2023). The making of momentum: A demand-system perspective. In *Proceedings of the EUROFIDAI-ESSEC Paris December Finance Meeting*.

Investment Company Institute (2025, March). Active and index investing, february 2025. `https://www.ici.org/research/stats/combined_active_index`. Accessed April 18, 2025.

Jansen, K. A. (2025). Long-term investors, demand shifts, and yields. *Review of Financial Studies 38*(1), 114–157.

Jansen, K. A., W. Li, and L. Schmid (2024). Granular treasury demand with arbitrageurs. Technical report, National Bureau of Economic Research.

Jiang, H., L. Zheng, and D. Vayanos (2025). Passive investing and the rise of mega-firms. *Review of Financial Studies forthcoming*.

Jiang, Z., R. J. Richmond, and T. Zhang (2022). Understanding the strength of the dollar. Technical report, National Bureau of Economic Research.

Jiang, Z., R. J. Richmond, and T. Zhang (2024). A portfolio approach to global imbalances. *The Journal of Finance 79*(3), 2025–2076.

Jones, C. (2002). A century of stock market liquidity and trading costs. *Available at SSRN 313681*.

Kandel, E. and N. D. Pearson (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy 103*(4), 831–872.

Kaul, A., V. Mehdrotra, and R. Morck (2000). Demand curves for stocks do slope down: New evidence from an index weights adjustment. *The Journal of Finance 55*(1), 893–912.

Koijen, R. S., F. Koulischer, B. Nguyen, and M. Yogo (2021). Inspecting the mechanism of quantitative easing in the euro area. *Journal of Financial Economics 140*(1), 1–20.

Koijen, R. S., R. J. Richmond, and M. Yogo (2024). Which investors matter for equity valuations and expected returns? *Review of Economic Studies 91*(4), 2387–2424.

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

Koijen, R. S. and M. Yogo (2020). Exchange rates and asset prices in a global demand system. Technical report, National Bureau of Economic Research.

Koijen, R. S. J. and M. Yogo (2025, May). On the Theory and Econometrics of (Demand System) Asset Pricing.

Kozak, S., S. Nagel, and S. Santosh (2018). Interpreting factor models. *The Journal of Finance 73*(3), 1183–1223.

Kvamvold, J. and S. Lindset (2018). Do dividend flows affect stock returns? *Journal of Financial Research 41*(1), 149–174.

Li, J. and Z. Lin (2022). Price multipliers are larger at more aggregate levels. *Available at SSRN 4038664*.

Lou, D. (2012). A flow-based explanation for return predictability. *The Review of Financial Studies 25*(12), 3457–3489.

Madhavan, A. (2003). The russell reconstitution effect. *Financial Analysts Journal 59*(4), 51–64.

Pástor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political economy 111*(3), 642–685.

Pavlova, A. and T. Sikorskaya (2022). Benchmarking intensity.

Petajisto, A. (2011). The index premium and its hidden cost for index funds. *Journal of empirical Finance 18*(2), 271–288.

Schmickler, S. (2020). Identifying the price impact of fire sales using high-frequency surprise mutual fund flows. *Available at SSRN 3488791*.

Schmickler, S. and P. Tremacoldi-Rossi (2022). Spillover effects of payouts on asset prices and real investment. *Available at SSRN 4287300*.

Shleifer, A. (1986). Do demand curves for stocks slope down? *The Journal of Finance 41*(3), 579–590.

Tamoni, A., S. Sokolinski, and Y. Li (2024). Which investors drive anomaly returns and how? *Available at SSRN 4242745*.

van der Beck, P. (2022). Short-versus long-run demand elasticities in asset pricing. *Asset Pricing (May 16, 2022). Swiss Finance Institute Research Paper* (22-67).

Vayanos, D. (1998). Transaction costs and asset prices: A dynamic equilibrium model. *Review of Financial Studies 11*, 1–58.

Wurgler, J. and E. Zhuravskaya (2002). Does arbitrage flatten demand curves for stocks? *The Journal of Business 75*(4), 583–608.

Appendix A Proofs

A.1 Proof of Theorem 1

We prove Theorem 1 under a weaker assumption than Assumption 1:

Assumption A.1. *Denote* $\beta^u_{i,S} \equiv \frac{\text{Cov}(u_{i,t},u_{S,t})}{\text{Var}(u_{S,t})}$ *as the coefficient of regressing the demand shift of investor* i*,* $u_{i,t}$*, on the aggregate demand shift,* $u_{S,t}$*. We have the following regularity condition:*

$$\widehat{\text{Var}}^{cs}_S(\frac{\zeta_i}{\zeta_S}) - 2\widehat{\text{Cov}}^{cs}_S(\frac{\zeta_i}{\zeta_S}, \beta^u_{i,S}) > 0$$

We first provide the proof of Theorem 1 under the relaxed assumption A.1, and discuss the intuition behind the condition.

Proof of Theorem 1. Given the demand curve equation (2) and price equation (4), we have

$$\Delta q_{i,t} = u_{i,t} - \frac{\zeta_i}{\zeta_S} u_{S,t},$$

Let $\sigma_i^2 \equiv \text{Var}(u_{i,t})$ denote the variance of investor i's demand shift, and $\sigma_{iS} \equiv \text{Cov}(u_{i,t}, u_{S,t})$ denote the covariance between investor i's demand shift and the aggregate demand shift. The variance of flows and price are:

$$\begin{aligned}
\sigma^2_{q,i} &= \text{Var}(\Delta q_{i,t}) = \sigma_i^2 - 2\frac{\zeta_i}{\zeta_S}\sigma_{iS} + \frac{\zeta_i^2}{\zeta_S^2}\text{Var}(u_{S,t}) \\
\sigma_p^2 &= \text{Var}(\Delta p_t) = \frac{1}{\zeta_S^2}\text{Var}(u_{S,t})
\end{aligned}$$

The size-weighted average flow volatility is:

$$\begin{aligned}
\sigma_q^2 &= \sum_i S_i \sigma^2_{q,i} \\
&= \hat{\mathbb{E}}^{cs}_S\left[\sigma_i^2\right] - 2\hat{\mathbb{E}}^{cs}_S\left[\frac{\zeta_i}{\zeta_S}\sigma_{iS}\right] + \hat{\mathbb{E}}^{cs}_S\left[\frac{\zeta_i^2}{\zeta_S^2}\right]\text{Var}(u_{S,t}) \\
&= \hat{\mathbb{E}}^{cs}_S\left[\sigma_i^2\right] - 2\left(\hat{\mathbb{E}}^{cs}_S\left[\frac{\zeta_i}{\zeta_S}\right]\hat{\mathbb{E}}^{cs}_S\left[\sigma_{iS}\right] + \widehat{\text{Cov}}^{cs}_S\left[\frac{\zeta_i}{\zeta_S}, \sigma_{iS}\right]\right) + \left(\hat{\mathbb{E}}^{cs}_S\left[\frac{\zeta_i}{\zeta_S}\right]^2 + \widehat{\text{Var}}^{cs}_S\left(\frac{\zeta_i}{\zeta_S}\right)\right)\text{Var}(u_{S,t})
\end{aligned}$$

Notice that:

$$\hat{\mathbb{E}}_S^{cs}\left[\frac{\zeta_i}{\zeta_S}\right] = \frac{1}{\zeta_S}\sum_i S_i\zeta_i = 1$$

$$\hat{\mathbb{E}}_S^{cs}\left[\sigma_{iS}\right] = \sum_i S_i \operatorname{Cov}\left(u_{i,t}, u_{S,t}\right) = \operatorname{Var}(u_{S,t})$$

The expression can be simplified as:

$$\sigma_q^2 = \hat{\mathbb{E}}_S^{cs}\left[\sigma_i^2\right] - \operatorname{Var}\left(u_{S,t}\right) - 2\widehat{\operatorname{Cov}}_S^{cs}\left[\frac{\zeta_i}{\zeta_S}, \sigma_{iS}\right] + \widehat{\operatorname{Var}}_S^{cs}\left(\frac{\zeta_i}{\zeta_S}\right)\operatorname{Var}\left(u_{S,t}\right)$$

Under Assumption A.1 that $\widehat{\operatorname{Var}}_S^{cs}(\frac{\zeta_i}{\zeta_S}) - 2\widehat{\operatorname{Cov}}_S^{cs}(\frac{\zeta_i}{\zeta_S}, \beta_{i,S}^u) > 0$, where we note that $\beta_{i,S}^u = \frac{\sigma_{iS}}{\operatorname{Var}(u_{S,t})}$, the condition becomes:

$$\widehat{\operatorname{Var}}_S^{cs}\left(\frac{\zeta_i}{\zeta_S}\right) - 2\widehat{\operatorname{Cov}}_S^{cs}\left[\frac{\zeta_i}{\zeta_S}, \frac{\sigma_{iS}}{\operatorname{Var}(u_{S,t})}\right] > 0$$

which implies $\widehat{\operatorname{Var}}_S^{cs}\left(\frac{\zeta_i}{\zeta_S}\right)\operatorname{Var}(u_{S,t}) - 2\widehat{\operatorname{Cov}}_S^{cs}\left[\frac{\zeta_i}{\zeta_S}, \sigma_{iS}\right] > 0$.

Therefore:

$$\sigma_q^2 \geq \hat{\mathbb{E}}_S^{cs}\left[\sigma_i^2\right] - \operatorname{Var}\left(u_{S,t}\right)$$

The ratio of σ_q^2 to σ_p^2 is given as:

$$\frac{\sigma_q^2}{\sigma_p^2} \geq \zeta_S^2\left(\frac{1}{\operatorname{Var}(u_{S,t})/\hat{\mathbb{E}}_S^{cs}\left[\sigma_i^2\right]} - 1\right).$$

Using the definition $\rho = \frac{\operatorname{Var}(u_{S,t})}{\hat{\mathbb{E}}_S^{cs}[\sigma_i^2]}$ from the main text, we get:

$$\frac{\sigma_q^2}{\sigma_p^2} \geq \zeta_S^2\left(\frac{1}{\rho} - 1\right).$$

Taking the square root and using $\mathcal{M} = \frac{1}{\zeta_S}$, we have the bound:

$$\mathcal{M} \geq \frac{\sigma_p}{\sigma_q} \times \sqrt{\frac{1}{\rho} - 1}$$

□

Remarks. As discussed in the main text, introducing heterogeneity in elasticities can increase the flow volatility for a given level of demand heterogeneity, as different responses to price changes provide another reason to trade other than heterogeneity in demand.

Assumption A.1 further relaxes the independence assumption in Assumption 1 by allowing for the cross-sectional dependence of data-generating processes on the demand shifts and the elasticity, captured by the cross-sectional covariance between elasticity and the correlation with the aggregate demand shift, $\widehat{\mathrm{Cov}}_S^{cs}(\frac{\zeta_i}{\zeta_S}, \beta_{i,S}^u)$.

The cross-sectional covariance between the elasticity and the correlation with aggregate shocks also affects the flow volatility. To see this more clearly, note that the condition in Assumption A.1 can also be expressed in terms of the correlation with the change in price:

$$\widehat{\mathrm{Cov}}_S^{cs}(\frac{\zeta_i}{\zeta_S}, \beta_{i,S}^u) = \frac{1}{\zeta_S^2}\widehat{\mathrm{Cov}}_S^{cs}(\zeta_i, \beta_{i,p}^u)$$

where $\beta_{i,p}^u = \frac{\mathrm{Cov}(u_{i,t}, \Delta p_t)}{\mathrm{Var}(\Delta p_t)}$ is the regression coefficient of the demand shift $u_{i,t}$ on the change in price Δp_t. Notice that though it is defined as the loading on the price, the causality runs the other way: demand shifts move the price, not vice versa.

All else equal, flow volatility can also be high because investors whose demand shifts move the price more (high $\beta_{i,p}^u$) are also less responsive to price changes ($\widehat{\mathrm{Cov}}_S^{cs}(\zeta_i, \beta_{i,p}^u) < 0$), and hence their demand shifts are more manifested in the observed trading. Empirically, large investors, who have larger weights in the aggregation and hence typically are more represented in the aggregate demand shifts, tend to be less responsive to price changes in proportion to their size relative to small investors, often due to trading costs or price impact concerns.

On the contrary, when investors whose demand shifts track the price closer are also more price-elastic, $\widehat{\mathrm{Cov}}_S^{cs}(\frac{\zeta_i}{\zeta_S}, \beta_{i,p}^u) > 0$, the opposite channel may dampen the observed flow volatility. Intuitively, their demand shifts are less passed through to the realized trading as they react to the disadvantageous price changes. When this force is overly strong, we may even end up in a pathological equilibrium where investors on average sell when they receive positive demand shifts.[23]

The condition in Assumption A.1 allows for the latter case, but essentially requires that it is dominated by the increase in flow volatility due to the dispersion in elasticities.

[23]To be precise, we may have the empirical moment such that $\hat{\mathbb{E}}_S^{cs}[\mathrm{Cov}(\Delta q_{i,t}, u_{i,t})] < 0$. Note that $\hat{\mathbb{E}}_S^{cs}[\mathrm{Cov}(\Delta q_{i,t}, u_{i,t})] = \hat{\mathbb{E}}_S^{cs}[\sigma_i^2 - \frac{\zeta_i}{\zeta_S}\mathrm{Cov}(u_{i,t}, u_{S,t})]$. Using the equality $\hat{\mathbb{E}}_S^{cs}[\frac{\zeta_i}{\zeta_S}\mathrm{Cov}(u_{i,t}, u_{S,t})] = \mathrm{Var}(u_S) + \widehat{\mathrm{Cov}}_S^{cs}\left[\frac{\zeta_i}{\zeta_S}, \sigma_{iS}\right]$ as in the proof, we can show that $\hat{\mathbb{E}}_S^{cs}[\mathrm{Cov}(\Delta q_{i,t}, u_{i,t})] < 0$ when $\widehat{\mathrm{Cov}}_S^{cs}(\frac{\zeta_i}{\zeta_S}, \beta_{i,S}^u) > \frac{1}{\rho} - 1$.

Appendix B Microfoundations

We start with demand curve representations of portfolio choice under CRRA utility in Section B.1. We then extend the analysis to a learning-from-price model in Section B.2.

B.1 CRRA Utility

Consider the portfolio choice problem of an investor with CRRA utility in a two-period model. With log-normal returns, the utility maximization gives the standard portfolio choice formula:

$$\frac{PQ_i}{W_i} = \frac{\mu - R_f}{\gamma_i \sigma_R^2}$$

where W_i is the investor's wealth, γ_i the risk aversion, $\mu \equiv \mathbb{E}\left[\frac{D}{P}\right]$ the expected return, R_f the risk-free rate, and σ_R the return volatility.

We perturb the portfolio-choice problem around a symmetric equilibrium where $\frac{PQ_i}{W_i} = 1$ with first-order log-linearization.[24] We use lowercase letters to denote the log of the uppercase counterpart, and use bar and Δ to indicate the symmetric equilibrium value and the deviation from the original equilibrium, respectively. We have:

$$\Delta q_i \approx -\underbrace{\frac{\bar{\mu}}{\bar{\mu} - \bar{r}}}_{\bar{\delta}} \Delta p + \underbrace{\frac{\bar{\mu}}{\bar{\mu} - \bar{r}} \mathbb{E}\left[\Delta d\right] - \Delta \log \gamma_i - \Delta \log \sigma_R^2}_{u_i} \tag{B.1}$$

In the CRRA model, the demand elasticity $\bar{\delta}$ is determined by the risk free rate and the expected return, which in the equilibrium is further pinned down by the return volatility and risk aversion; the demand shifter u_i comes from different sources, including changes in expectations about fundamentals ($\mathbb{E}\left[\Delta d\right]$), changes in risk aversion and uncertainty.

B.2 Learning-From-Price Model à la Hellwig (1980)

We extend the CRRA model to incorporate learning-from-price, adapting the framework from Hellwig (1980).

To focus on the learning-from-price mechanism, we consider a simplified version where demand shifts come only from heterogeneous expectations about dividend changes:

[24] By perturbing around the equilibrium with $\frac{PQ_i}{W_i} = 1$, we simplify the expression by eliminating the change in wealth on the left-hand side.

$$\Delta q_i = -\bar{\delta}\Delta p + \bar{\delta}\mathbb{E}_i[\Delta d] \tag{B.2}$$

The crucial assumption is that investors form expectations about dividend changes using both a private signal s_i and information extracted from the equilibrium price. We specify the information structure in detail later. Here, we postulate that expected dividend changes are formed as a linear combination:

$$\mathbb{E}_i[\Delta d] = \alpha_s s_i + \alpha_p \Delta p \tag{B.3}$$

where α_s and α_p are equilibrium coefficients that reflect how much weight investors place on their private signals versus price information.

Substituting (B.3) into (B.2) gives us the demand curve with learning-from-price:

$$\begin{aligned} \Delta q_i &= -\bar{\delta}\Delta p + \bar{\delta}\left(\alpha_s s_i + \alpha_p \Delta p\right) \\ &= -\underbrace{\bar{\delta}\left(1-\alpha_p\right)}_{\zeta}\Delta p + \underbrace{\bar{\delta}\alpha_s s_i}_{u_i} \end{aligned} \tag{B.4}$$

The key insight is that learning from prices makes demand less elastic: the effective elasticity $\zeta = \bar{\delta}(1-\alpha_p)$ is smaller than the elasticity $\bar{\delta}$ under rational expectations. When investors observe a price increase, they partly interpret it as conveying positive information about fundamentals, leading them to increase rather than decrease their demand.

To interpret the demand curve in the main text, Equation (B.4) is sufficient. For completeness, below we provide a full characterization of the equilibrium to pin down the coefficients α_s and α_p.

Equilibrium Characterization To fully characterize the equilibrium, we need to determine α_s and α_p. We consider a market with N agents of respective sizes S_i (where $\sum_i S_i = 1$), and eventually take N to infinity. We also consider noise traders who submit orders u_n.

Market clearing requires:

$$\sum_i S_i \Delta q_i = 0$$

From the demand equation (B.4), market clearing implies:

$$0 = -\zeta \Delta p + \bar{\delta}\alpha_s \sum_i S_i s_i + u_n$$
$$\Rightarrow \quad \Delta p = \frac{\bar{\delta}\alpha_s s_S + u_n}{\zeta} \tag{B.5}$$

where $s_S \equiv \sum_i S_i s_i$ is the size-weighted average signal. This can be rewritten as:

$$\Delta p = \frac{\alpha_s}{1-\alpha_p}\left(s_S + \underbrace{\frac{u_n}{\bar{\delta}\alpha_s}}_{\equiv s_N}\right)$$

where s_N represents the effective "noise signal" from noise trading.

Information Structure and Signal Extraction We assume the fundamental follows:

$$D = \bar{D}\exp\left(\Delta d - \frac{1}{2}\sigma^2_{\Delta d}\right)$$

where $\Delta d \sim \mathcal{N}(0, \sigma^2_{\Delta d})$.

Each investor receives a private signal $s_i \sim \mathcal{N}(0, \sigma_s^2)$ with correlation structure:

$$\text{Cov}(s_i, s_j) = \rho\sigma_s^2 \quad \text{for } i \neq j \tag{B.6}$$
$$\text{Cov}(s_i, \Delta d) = \beta\sigma_s^2 \tag{B.7}$$

In the limit as $N \to \infty$, the conditional expectation of Δd given signals s_i and the aggregate signal $s_S + s_N$ (a linear function of the price) is:

$$\mathbb{E}\left[\Delta d \mid s_i, s_S + s_N\right] = \frac{\beta\sigma_N^2}{\sigma_N^2 + \sigma_s^2\rho(1-\rho)}\, s_i + \frac{\beta\sigma_s^2(1-\rho)}{\sigma_N^2 + \sigma_s^2\rho(1-\rho)}\,(s_S + s_N) \tag{B.8}$$

where σ_N^2 is the variance of the noise signal s_N. The derivation is provided at the end of this section.

Using the price equation, we can express this conditional expectation in terms of s_i and Δp:

$$\mathbb{E}\left[\Delta d \mid s_i, \Delta p\right] = \alpha_s s_i + \alpha_p \Delta p$$

Matching coefficients, we obtain:

$$\alpha_s = \frac{\beta\sigma_N^2}{\sigma_N^2 + \sigma_s^2\rho(1-\rho)} \tag{B.9}$$

$$\alpha_p = \frac{\sigma_s^2(1-\rho)}{\sigma_N^2 + \sigma_s^2(1-\rho)} \tag{B.10}$$

Substituting back into the demand equation:

$$\Delta q_i = \underbrace{\bar{\delta}\frac{\beta\sigma_N^2}{\sigma_N^2 + \sigma_s^2\rho(1-\rho)}}_{u_i} s_i - \underbrace{\bar{\delta}\frac{\sigma_N^2}{\sigma_N^2 + \sigma_s^2(1-\rho)}}_{\zeta} \Delta p$$

The final elasticity expression reveals the trade-off inherent in learning from prices. On one hand, when private signals are less correlated across investors (low ρ), more new information can be extracted from the price, making the market more inelastic. On the other hand, when noise trader flows are larger (high σ_N^2), the price becomes a less precise signal, making the market more elastic.

Derivation of the conditional expectation formula

Proof. The signal covariance matrix is given as (treating each i as infinitesimally small):

$$\Sigma_s = \text{Var}\left(\begin{bmatrix} s_i \\ s_S + s_N \end{bmatrix}\right) = \begin{bmatrix} \sigma_s^2 & \rho\sigma_s^2 \\ \rho\sigma_s^2 & \rho\sigma_s^2 + \sigma_N^2 \end{bmatrix}$$

To compute the (2,2) entry, notice that:

$$\begin{aligned} \text{Var}(s_S) = \text{Var}\left(\sum_i S_i s_i\right) &= \sigma_s^2 \left(\sum_i S_i^2 + \sum_{i \neq j} S_i S_j \rho\right) \\ &= \sigma_s^2 \left((1-\rho)\sum_i S_i^2 + \rho \sum_i \sum_j S_i S_j\right) \\ &= \sigma_s^2 \left(\rho + (1-\rho)\mathcal{H}\right) \end{aligned}$$

where $\mathcal{H} = \sum_i S_i^2$. Taking the limit as $N \to \infty$, we have $\mathcal{H} \to 0$, so $\text{Var}(s_S) = \rho\sigma_s^2$.

The covariance between signals and Δd is:

$$\mathrm{Cov}(\Delta d, s_i) = \beta\sigma_s^2$$
$$\mathrm{Cov}(\Delta d, s_S + s_N) = \beta\sigma_s^2$$

Thus $\Sigma_{s,\Delta d} = \begin{bmatrix} \beta\sigma_s^2 \\ \beta\sigma_s^2 \end{bmatrix}$.

The conditional expectation is given by $\Sigma_{s,\Delta d}^T \Sigma_s^{-1} \begin{bmatrix} s_i \\ s_S + s_N \end{bmatrix}$. Computing this yields the formula in the main text. □

Appendix C Price Impact Bound with Substitution across Assets

When strong substitution exists across assets, the interpretation of the price impact bound requires extra care. To illustrate the point, in this section, we consider a classic arbitrage example: a single-name ETF e and its underlying stock s.

We start with the single-price representation of the demand system, as derived in the main text:

$$\begin{aligned}
\Delta q_{i,t}(s) &= -\underbrace{\zeta(s)\ (1 - \mathcal{Q}_{s\leftarrow e}\,\mathcal{Q}_{e\leftarrow s})}_{\tilde{\zeta}(s) = 1/\mathcal{M}(s)}\ \Delta p_t(s)\ +\ \underbrace{\mathcal{Q}_{s\leftarrow e}\, u_{S,t}(e)\ +\ u_{i,t}(s)}_{\tilde{u}_{i,t}(s)} \\
\Delta q_{i,t}(e) &= -\underbrace{\zeta(e)\ (1 - \mathcal{Q}_{e\leftarrow s}\,\mathcal{Q}_{s\leftarrow e})}_{\tilde{\zeta}(e) = 1/\mathcal{M}(e)}\ \Delta p_t(e)\ +\ \underbrace{\mathcal{Q}_{e\leftarrow s}\, u_{S,t}(s)\ +\ u_{i,t}(e)}_{\tilde{u}_{i,t}(e)}
\end{aligned} \tag{C.1}$$

To illustrate potential issues arising from strong cross-asset substitution, we consider the following numerical example.

A numerical example Consider an asymmetric market between the ETF and stock: the ETF is ten times smaller than the stock and less actively traded (small $\sigma_q(e)$), but the no-arbitrage condition ensures the ETF price closely tracks the stock.

Numerically, let the price volatilities be $\sigma_p(s) = \sigma_p(e) = 0.1$ with near-perfect correlation, $\mathrm{corr}(\Delta p_s, \Delta p_e) \approx 1.00$. Let the flow volatility be $\sigma_q(s) = 0.1$ for the stock, and $\sigma_q(e) = 0.01$ for the ETF. For convenience, assume investor agreement of $\rho(s) = \rho(e) = 0.5$, so $\sqrt{\frac{1}{\rho} - 1} = 1$. This gives aggregate demand shift volatilities $\sigma_{uS}(s) = 0.1$ and $\sigma_{uS}(e) = 0.01$, respectively.[25] For simplicity, assume demand shifts are uncorrelated across assets – introducing correlation does not change the main point.

The moments we specified above can be induced by the following demand elasticity matrix:

$$\Gamma = \begin{pmatrix} -\zeta(s) = -100.3 & \zeta(s,e) = 99.7 \\ \\ \zeta(e,s) = 998.5 & -\zeta(e) = -1002.5 \end{pmatrix}$$

Under this demand system, we have the following reduced-form moments:

$$\mathcal{M}(s) = 1, \quad \mathcal{M}(e) \approx 0.1, \quad \mathcal{Q}_{s\leftarrow e} \approx 0.099, \quad \mathcal{Q}_{e\leftarrow s} \approx 9.95, \quad \mathcal{A} = 1 - \mathcal{Q}_{s\leftarrow e}\,\mathcal{Q}_{e\leftarrow s} \approx 0.01$$

We choose the demand elasticity matrix Γ to be asymmetric, reflecting the size difference between markets: the stock market is ten times larger than the ETF market. Consequently, investor demand for the stock responds less to ETF price changes than vice versa. In terms of demand pass-throughs, a 1% demand shift to the ETF translates to approximately 0.1% effective demand shift to the stock, while a 1% demand shift to the stock generates an effective 10% demand shift to the ETF. This asymmetry illustrates when our bound remains appropriate despite strong substitution, and when it becomes less informative.

In this context, we interpret the key objects in the bound: the price impact $\mathcal{M}(s)$ and the investor agreement $\rho(s)$.

Interpreting price impact $\mathcal{M}$. As discussed in the main text, our bound correctly recovers the price impact $\mathcal{M}(\cdot)$. However, under strong substitution, price impact differs substantially from the reciprocal of price elasticity $\zeta(\cdot)$. In this example, the price impact for the stock is approximately 1, while the own-price elasticity is around 100.

Understanding this difference requires recognizing that elasticity is a partial-equilibrium concept holding all other prices constant, while price impact incorporates general-equilibrium effects. Elasticity captures demand response to stock price changes while *holding the ETF price constant.* Given the close arbitrage relationship, investors are highly sensitive to price discrepancies between the ETF and stock, leading to aggressive arbitrage trading. Conversely, price impact captures how stock prices respond

[25] Notice that by replacing the endogenous prices from the demand equation we again have $\Delta q_{i,t}(\cdot) = u_{i,t}(\cdot) - u_{S,t}(\cdot)$, hence $\sigma_q(\cdot) = \sigma_{uS}(\cdot)\sqrt{\frac{1}{\rho(\cdot)} - 1}$.

to demand shocks *accounting for general-equilibrium effects*: the demand shift also moves ETF prices, which recursively affects stock demand. The amplification factor $\mathcal{A} := (1 - \mathcal{Q}_{n\leftarrow n'}\,\mathcal{Q}_{n'\leftarrow n})$ illustrates this feedback loop: a 1% stock demand shift generates a $\mathcal{Q}_{n'\leftarrow n} \approx 9.95\%$ effective demand shift to the ETF, which feeds back to the stock as $\mathcal{Q}_{n\leftarrow n'}\mathcal{Q}_{n'\leftarrow n} \approx 0.99\%$.

Under very strong substitution, our bound identifies the price impact but becomes uninformative about the underlying elasticity. Which parameter matters depends on the research question. For understanding market-wide price responses, the reduced-form price impact is often the key quantity of interest – the reciprocal of the price impact can be loosely interpreted as the price elasticity of the combined stock-ETF system. However, if the goal is estimating arbitrage strength between the stock and ETFs, our bound is ill-suited. Such applications require cross-sectional identification strategies that compare differential asset responses to shocks, as developed in Chaudhary et al. (2023) and Haddad et al. (2025).

Interpreting investor agreement ρ. With multiple assets, the effective demand shift for one asset $\tilde{u}_{i,t}(s)$ depends not only on $u_{i,t}(s)$ but also on the aggregate demand shift to its substitute $u_{S,t}(e)$. This substitute term enters because it moves substitute prices, effectively shifting the demand curve for asset s through substitution effects. The coefficient for $u_{S,t}(e)$ is $\mathcal{Q}_{s\leftarrow e}$, justifying our interpretation as demand pass-through: it measures the effective demand shift to asset s from a unit aggregate demand shift to e.

Investor agreement ρ measures the comovement of the total demand shifts $\tilde{u}_{i,t}(s)$ across investors, which in this case has an additional common factor: the aggregate demand shifts to the substitute $u_{S,t}(e)$.

Under strong substitution, high investor agreement ρ can emerge even when investors strongly disagree on asset-specific fundamentals. Consider the ETF in our example: even though investors may have heterogeneous ETF demand shifts $u_{i,t}(e)$, they all recognize the close arbitrage relationship between the ETF and underlying stock. Since the underlying stock is a much larger market and much more actively traded than the ETF, stock fundamental shifts passed to the ETF $\mathcal{Q}_{e\leftarrow s}u_{S,t}(s)$ dominate ETF-specific demand shifts $u_{i,t}(e)$. In our numerical example, $\mathrm{Var}(\mathcal{Q}_{e\leftarrow s}u_{S,t}(s)) \approx 1$ while $\mathrm{Var}(u_{S,t}(e)) \approx 0.0001$.

As $u_{S,t}(s)$ is reflected in the stock price and shared across investors, this creates high agreement ρ. If an econometrician naively applies the bound to the ETF using $\rho = 0.5$ for asset-specific demand shifts, they would recover a counterfactually high price impact $\mathcal{M}(e) \approx \frac{0.1}{0.01}\sqrt{\frac{1}{0.5} - 1} \approx 10$. This overestimate occurs because the true agreement incorporating substitution effects is close to 0.9999. For ρ close to one, the bound is highly nonlinear and less informative.

Having close substitutes does not automatically invalidate the bound – it depends on the magnitude of demand shifts passed from substitutes. In our example, the bound remains informative for the stock: given the relative market size and activity, demand shifts originating from the ETF market are negligible ($\mathcal{Q}_{s\leftarrow e} \approx 0.1$ and $\sigma_{uS}(e) = 0.01$) compared to stock-specific demand shifts $\sigma_{uS}(s) \approx 0.1$. In this case, the investor agreement ρ is still mostly about the stock fundamentals. Using our bound, one can recover the price impact for the stock as $\mathcal{M}(s) = \frac{0.1}{0.1}\sqrt{\frac{1}{0.5} - 1} \approx 1$, close to the true value.

Appendix D Data Construction Details

D.1 Flow Measures

Quarterly trades $\Delta Q_{i,t}(n)$ and changes in shares outstanding $\Delta\bar{Q}_t(n) = \bar{Q}_t(n) - \bar{Q}_{t-1}(n)$ are adjusted for stock splits in quarter t. We construct trades by the residual investor as $\Delta Q_{0,t}(n) = \Delta\bar{Q}_t(n) - \sum_{i=1}^{I} \Delta Q_{i,t}(n)$. All results in the paper are robust to omitting the residual sector and constructing $\bar{Q}_t(n)$ (and the corresponding size weights) as the sum of institutional shares held. However, we prefer the construction of the residual sector as this effectively accounts for trades by the institutional sector as a whole, which would otherwise be omitted. Furthermore, scaling by institutional shares held leads to some large outliers for smaller stocks that are held by very few institutions. Quarterly trades in percent are denoted by $q_{i,t}(n) = \frac{\Delta Q_{i,t}(n)}{Q_{i,t-1}(n)}$. To reduce the effect of outliers, we also use the Davis-Haltiwanger growth rate (Davis and Haltiwanger, 1992), following Gabaix and Koijen (2021) $q_{i,t}(n) = \frac{2(Q_{i,t}(n) - Q_{i,t-1}(n))}{Q_{i,t}(n) + Q_{i,t-1}(n)}$. The results are robust to either definition. When using portfolio turnover as the $\mathcal{L}_1$ approximation of flow volatility (the size-weighted variance of $q_{i,t}(n)$), there is no need to express trades in percent, as portfolio turnover sums raw trades $\Delta Q_{i,t}(n)$ relative to supply. This makes portfolio turnover a more robust estimator, less sensitive to outliers, and the treatment of extensive versus intensive margin trades.

D.2 Portfolio Turnover at the Fund Level

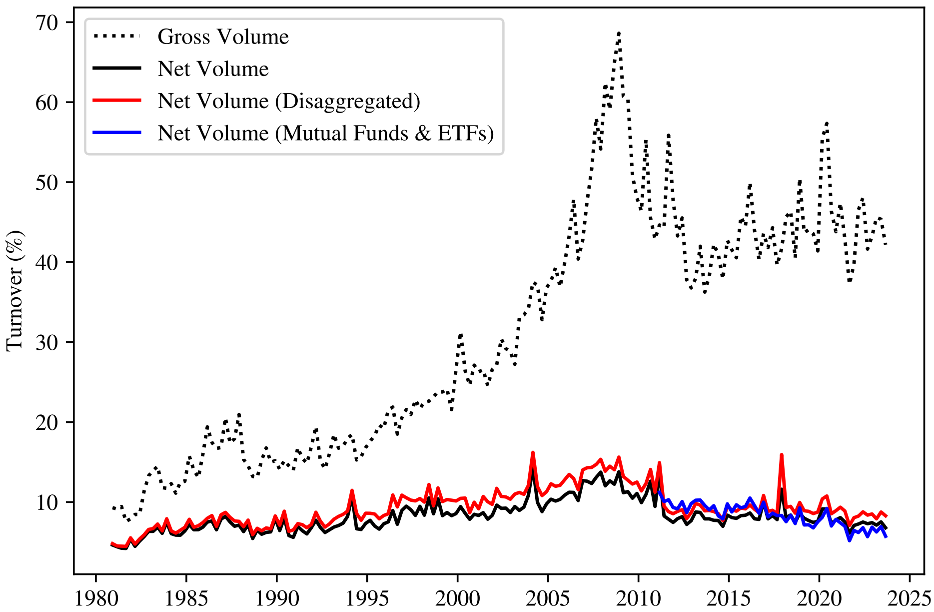
In the main text, we compute portfolio turnover at the 13F institution level to ensure comprehensive coverage. However, for asset managers with multiple subsidiary funds, institutional-level portfolio turnover excludes intrafamily transactions, which may potentially explain why portfolio turnover is smaller than gross volumes. This section uses disaggregated mutual fund holdings data to demonstrate that netting effects from within-institution aggregation are negligible.

We disaggregate fund families in the 13F institutional holdings data (S34 file) using Thomson Reuters mutual fund holdings data (S12 file). Using the S12-S34 link table, we match mutual fund holdings to their corresponding asset managers in the 13F data. For asset managers whose total holdings exceed the sum of their subsidiary fund holdings, we construct a residual entity representing the difference between institutional and mutual fund holdings. We retain institutions in the 13F data that are not matched to any mutual fund. We then compute portfolio turnover from this merged dataset using the same methodology as in the main text.

As an additional validation, we construct fund-level portfolio turnover using an alternative source: the CRSP Survivor-Bias-Free US Mutual Fund Database, which provides comprehensive coverage of mutual funds and ETFs but excludes other investor types. Since these funds account for a smaller share of market ownership than the broader 13F universe, we normalize portfolio turnover within the dataset – dividing net trading activity by the total shares held by all CRSP funds, rather than by shares outstanding. Normalizing by shares outstanding would yield much smaller portfolio turnover and render it noncomparable to that based on 13F data.

Figure D.1 compares portfolio turnover measures computed using these two approaches with our baseline institutional-level measures. The red line shows the portfolio turnover computed from the disaggregated 13F data using S12 mutual fund holdings files. The blue line presents portfolio turnover computed from the CRSP Survivor-Bias-Free US Mutual Fund Database. Despite being computed from different data sources and aggregation levels, the baseline institutional portfolio turnover is very close to the fund-level measures, confirming that netting effects from within-institution aggregation are negligible.

Figure D.1: **Portfolio Turnover at the Fund Level**
The figure compares portfolio turnover measures at the 13F institutional level with portfolio turnover computed at the fund level. *Portfolio Turnover* (in black) shows the baseline portfolio turnover computed from 13F institutional holdings data. *Portfolio Turnover Disaggregated* (in red) presents portfolio turnover computed from 13F data disaggregated using Thomson Reuters S12 mutual fund holdings files. *Portfolio Turnover Mutual Fund & ETFs* (in blue) presents portfolio turnover computed from the CRSP Survivor-Bias-Free US Mutual Fund Database, normalized within the dataset.



D.3 Measuring Investor Agreement from I/B/E/S

We measure investor agreement using analyst forecast data from I/B/E/S, leveraging the idea that the cross-sectional distribution of analyst beliefs serves as a proxy for the cross-sectional distribution of investor demand. This section details the sample construction and methodology for estimating the agreement parameter ρ.

D.3.1 Data Sources and Sample Selection

We obtain analyst earnings forecasts from the I/B/E/S Detail History database (`ibes.det_epsus`). We only use S&P 500 constituent firms to ensure sufficient number of forecasts. We then link I/B/E/S tickers to CRSP identifiers through a multi-step process: first matching I/B/E/S tickers to Compustat's `gvkey` using the security linking table (`comp.security`), then connecting `gvkey` to CRSP's `permno` through the CCM linking table using link types `LU` and `LC`. Finally, we filter for forecasts made while firms were S&P 500 constituents using historical index membership data.

We focus on two types of forecasts:

- **Quarterly Earnings-per-Share (EPS) forecasts** (FPI codes 6, 7, 8, 9): Representing 1-

through 4-quarter ahead EPS forecasts;

- **Long-term growth (LTG) forecasts** (FPI code 0): Representing long-term earnings growth rates.

D.3.2 Construction of Forecast Updates

We identify forecasters at the institution level (`estimator`, brokerage house or sell-side institution), to be consistent with the holdings data which is also at the 13F institution level.

For each forecaster-firm pair, we track how forecasts evolve over time:

EPS Forecast Updates: For quarterly EPS forecasts, we track how forecasters update their forecasts for a specific earnings announcement as it approaches. Each forecast target is uniquely identified by the firm and fiscal period end date (`fpedats`), with the actual earnings released on `anndats_act`. We define the forecast horizon as the number of days between when a forecast is made (`anndats`) and when actual earnings are released (`anndats_act`), converted to quarters by dividing by 90. We retain forecasts made within 400 days of the actual release and round horizons to the nearest quarter with a 30-day tolerance window. When multiple forecasts exist for the same forecaster-target-horizon combination, we select the earliest forecast.

Denote the forecasted EPS by forecaster i at time t for firm n and horizon h as $f_{i,t}^{h}(n)$. Updates are then calculated as percentage changes between consecutive horizons for the same target:

$$\Delta f_{i,t}^{h}(n) = f_{i,t}^{h}(n) - f_{i,t-1}^{h+1}(n)$$

By construction, $f_{i,t}^{h}(n)$ is around 90 days later than $f_{i,t-1}^{h+1}(n)$, matching the frequency of holdings data.

LTG Forecast Updates: Long-term growth forecasts differ from EPS forecasts as they lack a specific target date and thus no natural horizon measure. For these forecasts, we track quarter-to-quarter changes by assigning each forecast to a quarter based on its announcement date (`anndats`). To avoid partial quarter effects, forecasts made 45 or more days into a quarter are assigned to the following quarter. For each forecaster-firm-quarter combination, we retain only one forecast (the earliest if multiple exist). Updates are then calculated as simple differences (not percentages) between consecutive quarterly LTG forecasts:

$$\Delta f_{i,t}^{LTG}(n) = f_{i,t}^{LTG}(n) - f_{i,t-1}^{LTG}(n)$$

where $f_{i,t}^{LTG}(n)$ is the long-term growth forecast by forecaster i in quarter t for firm n.

D.3.3 Estimation of Agreement ρ

Following our theoretical framework, we estimate forecaster agreement $\rho(n)$ as the adjusted R^2 from regressing individual forecast updates on time fixed effects. Specifically, for each firm n and forecast type (EPS at horizon h or LTG), we then estimate:

$$\Delta \hat{f}_{i,t}^{h}(n) = \overline{\Delta f_t^h(n)} + \epsilon_{i,t}^{h}(n) \quad \text{for each } h \in \{1, 2, 3, LTG\}$$

where $\Delta \hat{f}_{i,t}^{h}(n) = \Delta f_{i,t}^{h}(n) - \overline{\Delta f_i^h(n)}$ are the demeaned forecast updates within each forecaster-horizon-firm combination, and $\overline{\Delta f_t^h(n)}$ are time fixed effects. The adjusted R^2 from this regression captures the proportion of forecast update variation explained by common time effects, serving as our measure of agreement $\rho_{EPS}^{h}(n)$.

We use adjusted R^2 as opposed to the original R^2, as the latter can incur a large bias when the number of forecasters is small: When there are only N forecasters, the expected raw R^2 will be around $\frac{1}{N}$ even with completely independent forecasts (hence the population R^2 is 0), while the adjusted R^2 have an expectation of 0 in this case. However, the adjusted R^2 can be negative in the sample. In these rare cases (mostly occur in the LTG forecasts when number of forecasters is small), we truncate the adjusted R^2 at 0.

To further reduce noises due to unbalanced panels, we apply the following filters before estimating $\rho_{EPS}^{h}(n)$: For each firm-horizon pair in quarterly EPS forecasts, we drop forecasters with less than 5 periods of forecast updates, and drop periods with less than 5 forecasters per firm-horizon combination. We repeat this filter iteratively until no more forecasters or periods can be dropped. The LTG forecasts are more sparse, hence we lower the threshold for the number of periods of forecast updates per forecaster-firm-quarter combination and the number of forecasters per firm-quarter combination to 4 and 3, respectively. Table D.1 reports the average characteristics of the final sample.

Table D.1: I/B/E/S Average Number of Forecasters and Updates
The table reports average characteristics of the I/B/E/S forecast sample used to estimate investor homogeneity. "N Periods" refers to the average number of time periods with forecasts per firm-horizon pair. "N Forecasters" is the average number of unique estimators covering each firm-horizon pair. "N Updates per Period" is the average number of forecast updates per firm-period. "N Updates" is the total average number of forecast updates per firm. 1Q-3Q refer to one-quarter through three-quarters ahead EPS forecasts, and LTG refers to long-term growth forecasts.

Horizon	**N Periods**	**N Forecasters**	**N Updates per Periods**	**N Updates**
1Q	44.2	26.0	11.5	510.2
2Q	41.3	25.0	10.9	452.4
3Q	37.6	22.6	10.1	379.8
LTG	16.5	5.2	3.5	57.7

D.4 Estimate Agreement $\rho(n)$ Structurally

We use a workhorse structural model, designed to jointly match portfolio holdings and prices, to infer stock-level ρ. To this end, we take Koijen and Yogo (2019), and estimate stock-level agreement ρ from the investor-level demand shifts implied by their model. We acknowledge that using the ρ computed in Koijen and Yogo (2019) for our bounds requires assuming that elasticities estimated from portfolio holdings in levels correspond to quarterly elasticities. However, van der Beck (2022) shows that such level-based estimates instead capture long-run elasticities—extending beyond a one-year horizon. The structurally inferred disagreement should therefore be interpreted only as suggestive evidence of disagreement. Koijen and Yogo (2019) propose the following logit demand curve each investor i and quarter t

$$\log \frac{w_{i,t}(n)}{w_{i,t(0)}} = \beta_{i,t}\text{me}_t(n) + X_t(n)\beta_{i,t} + \epsilon_{i,t}(n) \tag{D.1}$$

which can be microfounded from mean-variance optimal portfolio choice under specific coefficient constraints. $\text{me}_t(n) = \log \text{ME}_t(n)$ is the market cap of stock n and $X_t(n)$ includes the characteristics book equity, dividends-to-book equity, market beta, profitability, and investments. We estimate (D.1) via linear GMM using KY's mandate-based instrument and pooling investors with fewer than 1000 cross-sectional holdings by their assets under management. The moment condition is given by

$$\mathbb{E}_t[\epsilon_{i,t}(n)|\widehat{\text{me}}_{i,t}(n), X_t(n)] \quad \text{s.t.} \quad \beta_{i,t} < 1 \quad \forall i, t.$$

where $\widehat{\text{me}}_{i,t}(n)$ is the counterfactual log market equity obtained if every institution held an equal-weighted portfolio given its investment universe. For each investor, stock, and date, we extract the quarterly demand shifts $u_{i,t}(n)$ from

$$u_{i,t}(n) = \Delta \log Q_{i,t}(n) - (\hat{\beta}_{i,t} - 1)\Delta \text{me}_{i,t}(n) \tag{D.2}$$

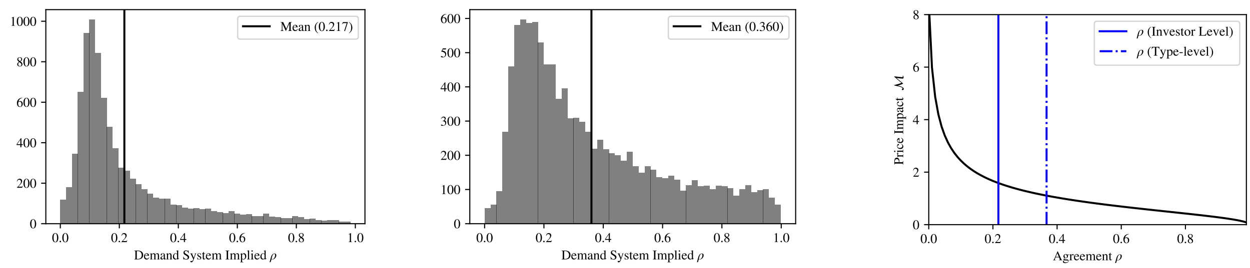
where $\Delta \log Q_{i,t}(n)$ and $\Delta\text{me}_{i,t}(n)$ are directly observable trades and changes in market equity, and $\hat{\beta}_{i,t}$ is the estimated coefficient on log market equity obtained from moment condition (D.4). Let $\tilde{u}_{i,t} = u_{i,t}(n) - \bar{u}_i(n)$ denote the time-series demeaned demand, and $S_{i,t}(n) = \frac{w_{i,t}(n)A_{i,t}}{\sum_i w_{i,t}(n)A_{i,t}}$ the investor-level size-weights for each stock. We then compute stock-level measure of investor agreement as

$$\rho(n) = 1 - \frac{\sum_t \sum_i S_{i,t-1}(\tilde{u}_{i,t} - \tilde{u}_{S,t})^2}{\sum_t \sum_i S_{i,t-1}\tilde{u}_{i,t}^2} \tag{D.3}$$

where $\tilde{u}_{S,t}$ is the size-weighted average of the time-series demeaned demand shifters $\tilde{u}_{i,t}$. Panel a) of Figure D.2 plots the $\rho(n)$ for each stock. Because investor-level $\beta_{i,t}$ are estimated with considerable noise, which may artificially inflate the cross-investor variation in $u_{i,t}(n)$ and therefore artificially deflate $\rho(n)$, we also compute group-level $\hat{\beta}_{i,t}$ by averaging across investors with the same 13F typecode. We then infer $u_{i,t}(n)$ by plugging in the averaged type-specific $\hat{\beta}_{i,t}$ and compute the corresponding ρ (Panel b). Panel c) plots the average ρ along with the average bound for the cross-section of US stocks.

Figure D.2: **Implied Stock-level Agreement $\rho(n)$ from KY (2019)**
Panel a) plots distribution of stock-level $\rho(n)$ implied by KY using investor-specific elasticities. Panel b) plots distribution of $\rho(n)$ under elasticities that vary by investor type. Panel c) plots the average ρ implied by KY along with the average bound.

(a) ρ from Investor-level $\beta_{i,t}$ (b) ρ from Group-level $\beta_{i,t}$ (c) Bound



D.5 Flow-Induced Trades by Mutual Funds

Our construction of flow-induced trades by mutual funds closely follows Lou (2012). We use quarterly mutual fund flows from the CRSP mutual fund survivorship-bias-free database. We set quarterly flows less than -100% or greater than 200% to missing and only include funds for which the total assets computed from their portfolio holdings are between 75% and 120% of the total net assets (TNA) reported by CRSP. Let $F_{i,t}$ denote the quarterly flow (in dollars) into fund i. Quarterly flow-induced demand is then simply given by summing over all hypothetical trades if flows were invested in line with previous portfolio weights $w_{i,t-1}(n)$:

$$FIT_t(n) = \frac{\sum_{i=1}^{I} F_{i,t} w_{i,t-1}(n)}{MV_{t-1}(n)} \tag{D.4}$$

where $MV_{t-1}(n)$ is the total market cap of stock n as of the previous quarter.

Appendix E Additional Figures and Tables

Figure E.1: **Portfolio Turnover versus Flow Volatility**

The figure plots the relationship between flow volatility $\sigma_q(n) = \sqrt{\sum_i S_i(n)\sigma_{q,i}^2(n)}$ and average portfolio turnover $\sqrt{\frac{\pi}{2}}\mathbb{E}[\frac{\sum_i |\Delta Q_i|}{Q^*}]$, which are size-weighted averages of $\mathcal{L}_2$ and $\mathcal{L}_1$ norms respectively.

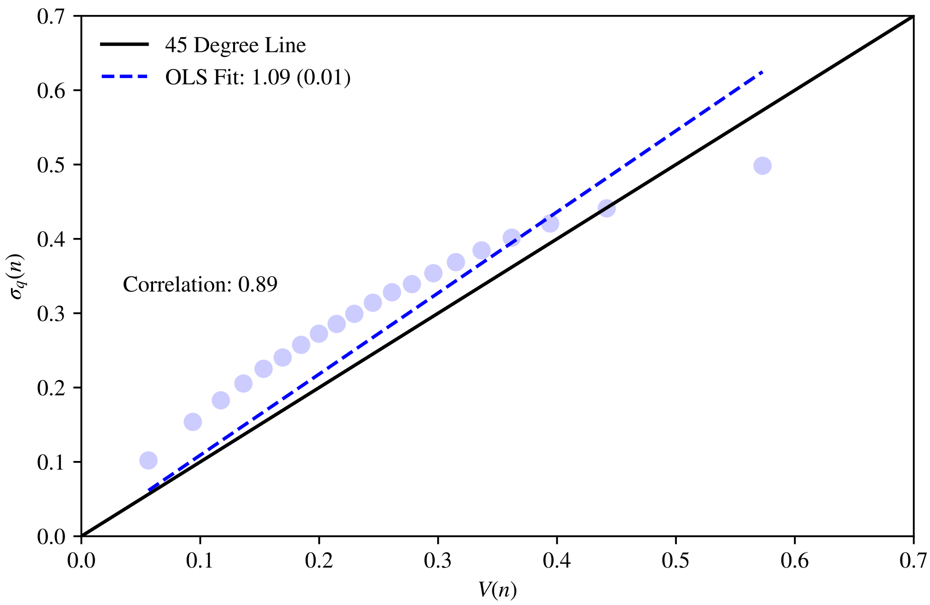


Figure E.2: **Empirical Relevance of ρ**
Panel a) plots the derivative $\frac{\partial \mathcal{M}}{\partial \rho} = -\frac{1}{2}\tilde{\mathcal{M}}\frac{1}{\rho^2}\sqrt{1/\rho - 1}$ as a function of ρ for the average US stock. Panel b) decomposes the variance of $\log \mathcal{M}_{\text{EPS}}$ into its underlying components $\log \sigma_p$, $\log \sigma_q$, $\log \mathcal{D}$ where $\mathcal{D} = \sqrt{1/\rho - 1}$.

(a) $\frac{\partial \mathcal{M}}{\partial \rho}$ for varying levels of ρ (b) Variance Decomposition of $\log \mathcal{M}_{\text{EPS}}$

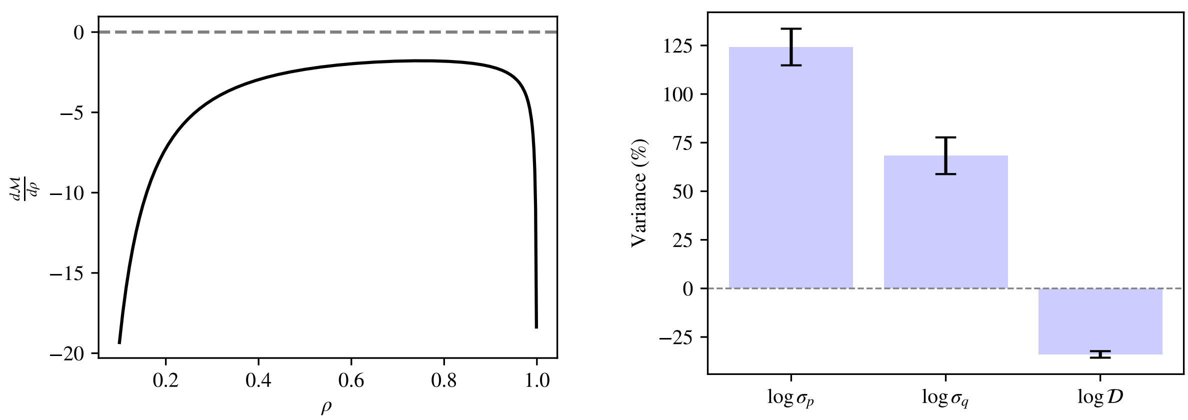


Table E.1: **Validation: Flow-Induced Trades**
The table summarizes the empirical validation of our bounds. We report the Panel coefficient of regressing quarterly stock-returns onto flow-induced trades (FIT) interacted with our bound $\mathcal{M}$, as well as the interaction with quintile dummies of $\mathcal{M}$. Formally, $r_t(n) = \alpha_t + \beta_1 FIT_t(n) + \beta_2 \mathcal{M}_t(n) + \beta_3\big(\mathcal{M}_t(n) \times FIT_t(n)\big) + \epsilon_t(n)$. T-stats are computed using standard errors clustered by date.

	Ret.		
	(1)	(2)	(3)
FIT	3.820***	2.635***	
	(0.510)	(0.636)	
$\mathcal{M}_{\text{EPS}}$		0.004	
		(0.004)	
FIT × $\mathcal{M}_{\text{EPS}}$		1.256*	
		(0.565)	
FIT × $\mathcal{M}_{\text{EPS}}$ quintile: 1			2.755***
			(0.624)
FIT × $\mathcal{M}_{\text{EPS}}$ quintile: 2			3.175***
			(0.588)
FIT × $\mathcal{M}_{\text{EPS}}$ quintile: 3			4.029***
			(0.577)
FIT × $\mathcal{M}_{\text{EPS}}$ quintile: 4			4.152***
			(0.718)
FIT × $\mathcal{M}_{\text{EPS}}$ quintile: 5			5.485***
			(0.938)
Date	x	x	x
$\mathcal{M}_{\text{EPS}}$ quintile	-	-	x
Observations	152862	152862	152862
R^2	0.249	0.250	0.250
R^2 Within	0.004	0.005	0.005

Significance levels: $*$ p $<$ 0.05, $**$ p $<$ 0.01, $***$ p $<$ 0.001. Format of coefficient cell: Coefficient (Std. Error)

Table E.2: **Validation: S&P500 Inclusions**

The table summarizes the price impact of index inclusions and their relationship with our bound. We report the coefficient of regressing (signed) abnormal event returns during S&P500 index reconstitutions onto the bound $\mathcal{M}$. T-stats are computed using standard errors clustered by date. Columns (1)–(3) report results for the full sample period, while columns (4)–(6) restrict the analysis to the pre-2000 subsample.

	Abnormal Return					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathcal{M}_{\text{EPS}}$		0.058* (0.025)			0.110* (0.046)	
$\mathcal{M}_{\text{EPS}}$ quintile: 1			0.046*** (0.014)			-0.015 (0.028)
$\mathcal{M}_{\text{EPS}}$ quintile: 2			0.082*** (0.024)			0.058 (0.032)
$\mathcal{M}_{\text{EPS}}$ quintile: 3			0.073*** (0.016)			0.082** (0.028)
$\mathcal{M}_{\text{EPS}}$ quintile: 4			0.076*** (0.016)			0.082** (0.029)
$\mathcal{M}_{\text{EPS}}$ quintile: 5			0.140*** (0.024)			0.174*** (0.039)
Intercept	0.080*** (0.008)	0.039 (0.021)		0.088*** (0.016)	-0.002 (0.039)	
Observations	837	686	686	390	239	239
R^2	0.021	0.036	0.040	0.038	0.091	0.096
Adj. R^2	0.017	0.029	0.027	0.028	0.072	0.061

Significance levels: $*$ p $<$ 0.05, $**$ p $<$ 0.01, $***$ p $<$ 0.001. Format of coefficient cell: Coefficient (Std. Error)

Table E.3: **Flow-Induced Trading: Alternative Impact Measures**

The table summarizes the coefficient of regressing quarterly stock-returns onto flow-induced trades (FIT) interacted with our bound $\mathcal{M}$, the simplified bound $\tilde{\mathcal{M}}$, the bound constructed from CRSP total volume $\frac{\sigma_p}{\text{CRSP Vol.}}$, as well as Amihud liquidity. T-stats are computed using standard errors clustered by date.

	Ret.			
	$\mathcal{M}_{\text{EPS}}$ (1)	$\tilde{\mathcal{M}}$ (2)	$\frac{\sigma_p}{\text{CRSP Vol.}}$ (3)	Amihud Illiquidity (4)
FIT	2.635*** (0.636)	2.505*** (0.622)	3.376*** (0.574)	3.772*** (0.568)
$\mathcal{M}$	0.004 (0.004)	0.004 (0.005)	0.023 (0.012)	-0.001*** (0.000)
FIT × $\mathcal{M}$	1.256* (0.565)	1.519* (0.623)	2.352 (1.277)	0.014 (0.050)
Date	x	x	x	x
Observations	152862	152862	152862	152862
R^2	0.250	0.250	0.250	0.250
R^2 Within	0.005	0.005	0.005	0.005

Significance levels: $*$ p $<$ 0.05, $**$ p $<$ 0.01, $***$ p $<$ 0.001. Format of coefficient cell: Coefficient (Std. Error)

Table E.4: **S&P500 Inclusions: Alternative Impact Measures**
The table summarizes the price impact of index inclusions and their relationship with our bound. We report the coefficient of regressing (signed) abnormal event returns during S&P500 index reconstitutions onto the bound $\mathcal{M}$, the simplified bound $\tilde{\mathcal{M}}$, the bound constructed from CRSP total volume $\frac{\sigma_p}{\text{CRSP Vol.}}$. T-stats are computed using standard errors clustered by date.

	Abnormal Return			
	$\mathcal{M}_{\text{EPS}}$	$\tilde{\mathcal{M}}$	$\frac{\sigma_p}{\text{CRSP Vol.}}$	Amihud Illiquidity
	(1)	(2)	(3)	(4)
$\mathcal{M}$	0.058*	0.114***	0.062	0.008
	(0.025)	(0.027)	(0.046)	(0.004)
log(ME)	-0.000	0.008	0.007	0.003
	(0.013)	(0.011)	(0.013)	(0.014)
β	0.009	0.007	0.018	0.020
	(0.011)	(0.010)	(0.010)	(0.010)
$\frac{\text{Dividend}}{\text{BE}}$	0.011	0.005	0.005	0.008
	(0.008)	(0.007)	(0.007)	(0.008)
Profit	-0.021*	-0.023*	-0.025**	-0.021*
	(0.010)	(0.009)	(0.009)	(0.010)
Observations	686	837	837	666
R^2	0.036	0.049	0.025	0.027
Adj. R^2	0.029	0.043	0.019	0.020

Significance levels: $*$ p $<$ 0.05, $**$ p $<$ 0.01, $***$ p $<$ 0.001. Format of coefficient cell: Coefficient (Std. Error)

Figure E.3: **Price Impact at Different Horizons**
The figure plots the stock-level p/q volatility ratio $\tilde{\mathcal{M}}$ for the average stock from 1990 to 2024. $\tilde{\mathcal{M}}$ uses portfolio turnover and return volatility at the daily, quarterly, and annual frequency.

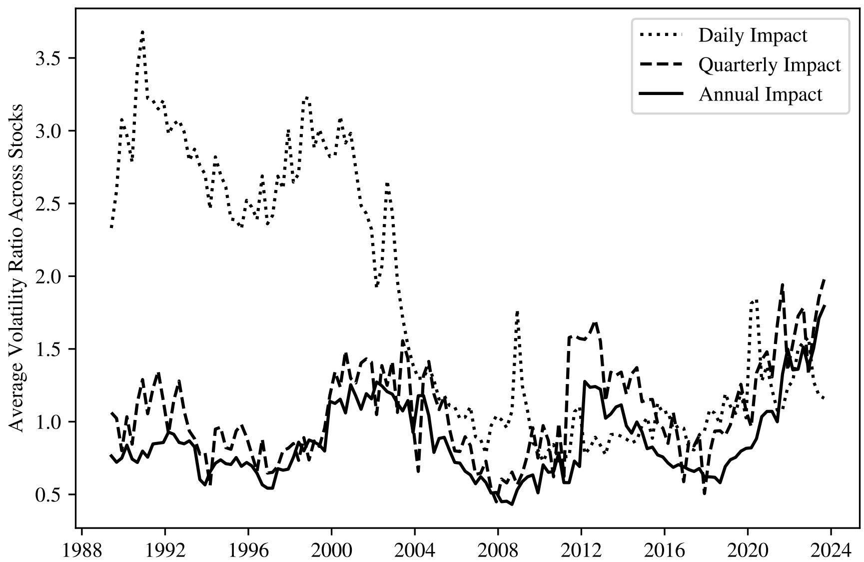


Table E.5: **Monthly Fama-MacBeth Regressions of Stock Returns**

This Table presents monthly Fama-MacBeth regressions of stock returns on price impact bound measures. The dependent variable is the monthly four-factor-adjusted return of month t, based on the four factors from Ken French's website (MKT, SMB, HML, and UMD), where the factor loadings are estimated in the preceding 60 months. The independent variables are the natural logs of our five empirical measures of the price impact bound. We estimate a cross-sectional regression in each month and then report the time series means and the t-statistics (in parentheses). We also report the time-series averages of the number of observations and adjusted R^2 of the cross-sectional regressions. All regressions include a constant, which is not reported for brevity. t-statistics are calculated using Newey-West robust standard errors with six lags.

	Dependent Variable: Four-Factor-adjusted Returns				
$ln(\tilde{M})$	0.340***				
	(4.51)				
$ln(M_{LTG})$		0.281***			
		(3.77)			
$ln(M_{1Q\ EPS})$			0.292***		
			(3.92)		
$ln(M_{2Q\ EPS})$				0.269***	
				(3.49)	
$ln(M_{3Q\ EPS})$					0.281***
					(3.65)
Adj. R^2	0.004	0.003	0.003	0.003	0.003
Avg. Obs	996	1023	1023	1023	1023
# Month	501	428	428	428	428

Significance levels: $*$ p $<$ 0.1, $**$ p $<$ 0.05, $***$ p $<$ 0.01.