Institutional Ownership Concentration and Informational Efficiency

Yan Xiong, Liyan Yang, and Zexin Zheng*

01/08/26

Abstract

This work studies how the concentration of ownership among institutional investors influences price informativeness in financial markets. We find that an increase in institutional ownership concentration — whether measured by investors' assets under management (AUM) or their firm-level holdings —reduces price informativeness as well as investment-to-price sensitivity. This negative effect is attributed to the learning and trading decisions of active, rather than passive, investors. To establish causality, we utilize a setting involving mergers between active investors, and our results remain consistent across both U.S. and international contexts.

Keywords: Ownership concentration, market power, active trading, price informativeness

JEL Classification: G14, G15, G23

*Xiong: University of Hong Kong; E-mail: yanxiong@hku.hk. Yang: Rotman School of Management, University of Toronto. E-mail: liyan.yang@rotman.utoronto.ca. Zheng: Hong Kong University of Science and Technology; E-mail: zzhengaq@connect.ust.hk . We thank Ferhat Akbas, Utpal Bhattacharya, Emilio Bisetti, Alon Brav, Maria Chaderina, Zhanhui Chen, Jaewon Choi, Eduardo Dávila, Matthew Dearth, Itay Goldstein, Ahmed Guecioueur, Jungsuk Han, Byoung-Hyoun Hwang, Marcin Kacperczyk, John Kuong, Pete Kyle, Jiyoon Lee, Jongsub Lee, Kuan-Hui Lee, Siguang Li, Yufei Lin, Baixiao Liu, Roger Loh, Fangyuan Ma, Linlin Ma, Andrey Malenko, Abhiroop Mukherjee, Don Noh, Seungjoon Oh, Haerang Park, Lubomir Petrasek, Sungjune Pyun, Fabrice Riva, Anna Scherbina, David Schoenherr, Tao Shen, Christof Stahel, Laura Starks, Kaushik Vasudevan, Tianyu Wang, John Wei, Jun Yang, Elaine Zhang, Leifu Zhang, Yi Zhang, Shan Zhao, and participants at 2024 Annual Conference on Asia-Pacific Financial Markets (CAFM), 2024 Greater Bay Area Finance Workshop, 2024 PHBS-CUHKSZ Economics and Finance Workshop, 2024 RUC Greater Bay Area Finance and Economics Alumni Academic Conference, 2025 ABFER Annual Conference Poster Session, 2025 China International Conference in Finance (CICF), 2025 EUROFIDAI-ESSEC Paris December Finance Meeting, 2025 European Accounting Association (EAA) Annual Congress, 2025 Financial Intermediation Research Society (FIRS) Conference, 2025 Fixed Income and Financial Institutions (FIFI) Conference, 2025 FTG Summer School Evening Session, 2025 Melbourne Asset Pricing Meeting, 2025 Northern Finance Association (NFA) Annual Meeting, 2025 SIEM Big Data and Market Microstructure Conference, and seminars at HKUST, HKUST(GZ), Nanyang Technological University, Seoul National University, and Yonsei University.

1 Introduction

Over the past several decades, the market share of U.S. industries has become increasingly concentrated amongst a select few (e.g., Autor et al., 2020; Kwon et al., 2024). The asset management industry is no exception. As illustrated in Figure 1, the concentration of institutional investors within the asset management space has been on the rise since 1980, with a particularly pronounced increase observed amongst active institutional investors. Beyond the more conspicuous concerns for economic and social equity, disproportionate ownership may also have reverberating effects on financial markets.

In this paper, we study polarized ownership structures and their implications for financial markets, specifically price informativeness, an established measure in the literature with important impacts on welfare and the real economy (e.g., Bond et al., 2012; Goldstein, 2023).

To guide our empirical analysis, we present a model based on Kacperczyk et al. (2025) and formalize the relationship between institutional ownership concentration and price informativeness. The model features heterogeneous assets and investors. Multiple risky assets of varying sizes are traded in the financial market, which consists of two types of traders: atomless competitive traders (e.g., retail investors) who take prices as given when trading and oligopolistic institutional investors who recognize that their trades can move asset prices. Such institutional investors differ in size, which determines the magnitude of their price impact, and can be classified as active or passive. Active investors have the capacity to collect information and thus reduce uncertainties regarding asset payoffs when trading; passive and retail investors lack information-gathering capabilities. After active investors make their learning choices, all traders trade in the financial market.

Our theoretical contribution is not a new model framework, per se; rather, we reinforce and then extend Kacperczyk et al.'s (2025) model to conduct two novel analyses, which inform our empirical analysis. First, we propose and explore a new measure of institutional ownership concentration that centers on investors' asset demand. In Kacperczyk et al. (2025), ownership concentration is measured by the primitive size distribution among institutional investors, which

we refer to as *size*-based ownership concentration. This can be empirically proxied using institutional investors' assets under management (AUM). Our new measure, termed *trading*-based ownership concentration, focuses on the asset demand of institutional investors. This measure is closely connected to investors' firm-level equity holding data, allowing us to leverage this rich dataset to enhance the power of our subsequent empirical analysis.

We find that, as the size of large institutional investors increases while that of small institutional investors decreases (i.e., indicating higher size-based ownership concentration), large investors increase demand for risky assets whereas small investors pursue the opposite—a pattern that suggests higher trading-based ownership concentration. In this way, Kacperczyk et al.'s (2025) measure is positively related to ours. Together with their finding that increased size-based ownership concentration leads to lower price informativeness, we conclude that increased trading-based ownership concentration is also negatively related to price informativeness.

Second, compared to Kacperczyk et al. (2025), our examination explicitly differentiates between ownership concentration amongst active and passive investors. While passive investors do not engage in information gathering or informed trading and thus do not directly impact price informativeness, changes in their ownership concentration could indirectly affect price informativeness by influencing the trading decisions of active investors. Our numerical analysis, however, indicates that it is an increase in ownership concentration among active, rather than passive, investors that suppresses price informativeness; as such, our empirical analysis focuses on ownership concentration among active investors.

We first investigate the impact of size-based active ownership concentration, which we quantify using two metrics: the Herfindahl-Hirschman Index (HHI) of AUM among active institutional investors and the share of AUM held by the top five active institutional investors. Across various model specifications, we observe a statistically significant and economically meaningful negative correlation between size-based active ownership concentration and price informativeness, thus affirming Kacperczyk et al.'s (2025) predictions. For instance, a 1%-point increase in active ownership concentration is associated with a 25.7% decrease in price informativeness relative

to its mean level. Furthermore, we explore whether this diminished informativeness affects real firm decisions. Our findings indicate that as active ownership concentration increases and suppresses price informativeness, the capacity for current prices to predict future firm investment also declines.

Despite their significance and robustness, the results derived from size-based ownership concentration can be limited by the small sample size. Therefore, we shift our focus to the new trading-based ownership concentration measure. This measure is defined similarly to the size-based one but is based on investors' holdings in each stock. We examine the relationship between active ownership concentration and price informativeness at the stock level. Our findings reveal that the price informativeness of stocks with the highest active ownership concentration is significantly lower than that of stocks with the lowest concentration. This effect is both statistically and economically significant across short and long prediction horizons. We also observe a decline in the predictive power of current stock prices for future investments.

Meanwhile, the loadings of passive ownership concentration in our specifications are unstable, varying from positive to negative, significant to insignificant. This instability indicates that ownership concentration amongst passive investors again does not drive our results, which aligns with our model's implications.

Given the endogenous ownership structure, the above regression results may be difficult to interpret economically. To address this concern, we leverage a quasi-natural experiment involving financial institution mergers. Specifically, the merger of two active institutional investors can lead to a plausibly exogenous increase in the active ownership concentration of any stocks held by both the acquirer and the target financial institutions. We find that, for these stocks, the subsequent decrease in price informativeness is significantly greater relative to other stocks held by one of the two merging parties.

Our findings are further solidified by a series of robustness tests. Notably, the negative relationship between active ownership concentration and price informativeness persists under alternative common measures of price informativeness, including post-earnings-announcement drift

(Ball and Brown, 1968), conditional probability of an information event (Duarte et al., 2020), informed trading intensity (Bogousslavsky et al., 2024), variance ratio (Lo and MacKinlay, 1988), and relative price informativeness (Dávila and Parlatore, 2025). Moreover, our results hold within an international context, extending beyond the U.S. market.

Finally, we investigate how active ownership concentration undermines price informativeness via two mechanisms introduced by Kacperczyk et al. (2025): the information pass-through effect and the learning effect. First, the information pass-through effect suggests that, as active institutional ownership becomes more concentrated, large active investors adopt conservative trading strategies to minimize their price impact. This channel is supported by empirical evidence of lower portfolio turnovers occurring when investors hold large shares in these portfolio companies. In addition, information content is reduced in earnings announcements for stocks with concentrated active institutional ownership. Second, the learning effect posits that the polarization of investor sizes hampers small investors' abilities to diversify their learning, leading them to focus on assets with the largest supply. Using downloads from the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system as a proxy for information acquisition, we empirically show that, as ownership concentration rises, EDGAR downloads between large and small stocks become increasingly imbalanced. Together, the information pass-through and learning effects contribute to the negative impact of active ownership concentration on price informativeness.

Our paper contributes to the literature on ownership structure. It relates most closely to Kacperczyk et al. (2025), who developed a general equilibrium model to analyze the joint impact of the size, concentration, and active and passive ownership shares of large investors on price informativeness. In addition to providing the first empirical evidence for size-based ownership concentration's predicted negative effect, we extend Kacperczyk et al.'s (2025) analysis by introducing a new trading-based ownership concentration measure, which we predict also negatively affects price informativeness. More recently, Glebkin et al. (2025) constructed a non-Constant Absolute Risk Aversion (CARA)-normal model to examine how asset concentration among a few large investors affects asset prices and liquidity; however, their model abstracts from asymmetric

information among investors.

Various empirical research have examined the implications of ownership concentration on financial markets. Greenwood and Thesmar (2011) find that stocks with concentrated ownership exhibit increased fragility, being more vulnerable to non-fundamental risk as indicated by stock return volatility. Managerial expectations of potential future misvaluation caused by this price fragility consequently lead to elevated precautionary cash holdings and reduced investment (Friberg et al., 2024). Porras Prado et al. (2016) demonstrate that ownership concentration results in increased short-selling restrictions due to blockholders who, fearing a loss of monitoring control, become reluctant to lend shares; this creates supply-side barriers that impede arbitrageurs from correcting mispricings, thereby inhibiting the injection of negative information. Massa et al. (2021) analyze the effects of an anticipated increase in ownership concentration following the merger of BlackRock and Barclays Global Investors. They report that the expected rise in concentration prompts shareholders to sell, leading to negative impacts on both price levels and liquidity. Huang et al. (2024) demonstrate that, in the corporate bond market, higher mutual fund ownership concentration leads to increased bond volatility. Kacperczyk et al. (2005) and Ivković et al. (2008) show that investors with informational advantages tend to hold concentrated portfolios, which can encourage information acquisition regarding illiquid, information-sensitive assets (Sialm and Xu, 2023).

Differing from the generalized focus on firm-level ownership concentration irrespective of institution size, Ben-David et al. (2021) investigate the implications of ownership concentration among the top ten largest institutional investors at the market level. They show that large institutions are granular in the sense that capital flows and trading strategies are correlated across different entities within the same institution; such granularity reduces the diversification of idiosyncratic shocks and induces higher price volatility. Meanwhile, Cen et al. (2024) divert their attention away from ownership concentration and examine the concentration of data scientists across all institutional investors holding a stock. They find that this concentration negatively affects price informativeness via the competition effect.

We build upon these prior works by exploring the effect of ownership concentration, particularly among active institutional investors, on price informativeness and investment-to-price sensitivity. Most importantly, we document a distinct mechanism through the learning and trading decisions of active investors, which is distinct from other drivers such as the exposure to non-fundamental demand shocks (Greenwood and Thesmar, 2011; Friberg et al., 2024; Huang et al., 2024), the role of short selling (Porras Prado et al., 2016), investors' responses to anticipated changes in ownership concentration (Massa et al., 2021), or the traits of granularity embedded in large institutions (Ben-David et al., 2021).

Other features of ownership structure have also been studied for their effect on price informativeness, including the total size of institutional ownership (Boehmer and Kelley, 2009), passive ownership (Bennett et al., 2020; Coles et al., 2022; Sammon, 2024), short-term ownership (Yan and Zhang, 2009), hedge fund ownership (Cao et al., 2018), socially responsible institutional ownership (Cao et al., 2023), and active funds' persistent aggregate capital flows (Dong et al., 2025). Our research distinguishes itself by focusing on ownership concentration and the welfare-based measure of price informativeness (Bai et al., 2016; Kacperczyk et al., 2021), which assesses the predictability of future earnings from current market prices.[1] Unlike commonly used price-based efficiency measures, this approach aligns closely with our theoretical framework and facilitates the examination of price informativeness's real effect on investment decisions (Bond et al., 2012; Goldstein, 2023).

The rest of the paper is organized as follows: Section 2 presents a theoretical framework and hypotheses to guide ensuing empirical investigations. Section 3 describes the data. Sections 4 and 5 present our main empirical findings using size- and trading-based ownership concentration measures, respectively. Section 6 delves into an analysis of the underlying mechanism. Section 7 concludes.

[1]There are also studies that examine the implications of product market concentration for stock market efficiency (e.g., Peress, 2010).

2 A Model of Ownership Concentration and Price Informativeness

This section presents a theoretical model based on Kacperczyk et al. (2025) to formalize the relationship between institutional ownership concentration and price informativeness. Our contribution is the novel analysis conducted within the economy established by Kacperczyk et al. (2025) that more effectively informs our empirical analysis.

First, we develop and explore the new *trading*-based ownership concentration measure, which differentiates itself from Kacperczyk et al.'s (2025) *size*-based measure defined by harnessing the granular firm-level investor holding data utilized in our empirical analysis. Second, to further explore the underlying mechanisms, we distinguish ownership concentration between active and passive investors and analyze their respective effects on price informativeness.

2.1 Setup

The model features a unit continuum of investors and multiple assets. There are three dates: $t = 0, 1,$ and 2. On date 0, investors acquire private information about the assets. On date 1, investors trade in the asset markets. On date 2, the assets pay off, and all investors consume.

Assets. There are one risk-free asset and $n \geq 1$ risky assets. The price of the risk-free asset is normalized to 1, its gross return is r, and it is in unlimited supply. Each risky asset is traded at an endogenous price $\tilde{p}_i$ per unit on date 1, and it pays an uncertain cash flow $\tilde{z}_i \sim N(\bar{z}, \sigma_i^2)$ on date 2, with $\bar{z} > 0$ and $\sigma_i > 0$. The total supply of risky asset i is $\tilde{x}_i \sim N(\bar{x}_i, \sigma_{xi}^2)$, with $\bar{x}_i > 0$ and $\sigma_{xi} > 0$, which is independent across assets and of all other random variables in the model.

Investors and trading. There is a continuum of investors, indexed by $j \in [0, 1]$. Investors allocate information capacity across assets on date 0, trade assets on date 1, and consume on date 2. All investors derive expected utility over their date 2 wealth according to a mean-variance

utility with a common risk-aversion coefficient $\rho > 0$. Without loss of generality, we normalize investors' initial wealth to zero.

The continuum of investors is divided into two groups, retail and institutional investors, who differ in their price impacts. Specifically, a mass $\lambda_0 < 1$ of investors are competitive atomistic uninformed investors (i.e., retail investors) indexed collectively by $j = 0$. Others are oligopolists, i.e., institutional investors who behave strategically and have price impacts. There are a number l of oligopolistic investors, each of whom, indexed by $j \in \{1, ..., l\}$, has information-gathering capacity K_j and size λ_j such that $\sum_{j=1}^{l} \lambda_j = 1 - \lambda_0$. Parameterized by λs, the sizes of the oligopolistic investors are mapped monotonically to ownership, serving as proxies for ownership shares, which in turn influence the investors' price impacts.

Per Kacperczyk et al. (2025), within the institutional category, investors are classified into the following four distinct types, based on their size and information-gathering capacity: (i) Large active investors (LA) who are large in size λ_j and capacity K_j; (ii) large passive investors (LP) who are large in size but have zero capacity; (iii) small active investors (SA) who are small in size and lower in capacity compared to large active investors; and (iv) small passive investors (SP) who are small in size and have zero capacity.

Prior to trading, on date 0, active investors ($j \in SA \cup LA$), who have positive information-gathering capacities, can acquire private signals regarding risky asset payoffs. Investor j's signal about the asset fundamental $\tilde{z}_i$ takes the form $\tilde{z}_i = \tilde{s}_{ji} + \tilde{\delta}_{ji}$, where $\tilde{\delta}_{ji}$ represents information loss due to constraints placed on learning and is independent of the signal $\tilde{s}_{ji}$. For retail and passive investors, $\tilde{s}_{ji} = \bar{z}$.

The vector of asset fundamental is denoted as $\tilde{z} = (\tilde{z}_1, ..., \tilde{z}_n)$, the vector of asset prices is denoted as $\tilde{p} = (\tilde{p}_1, ..., \tilde{p}_n)$, and investor j's private signal about the assets is denoted as $\tilde{s}_j = (\tilde{s}_{j1}, ..., \tilde{s}_{jn})$. Investor j's information set is $\mathcal{F}_j$. Following Kacperczyk et al. (2025), active and retail investors learn from prices whereas passive investors do not.[2] Since only active investors can acquire private signals regarding the assets, $\mathcal{F}_j = \{\tilde{p}, \bar{z}\}$ for $j = 0$, $\mathcal{F}_j = \{\tilde{p}, \tilde{s}_j\}$ for $j \in$

[2] As in Kacperczyk et al. (2025), our model implications remain robust whether passive investors are allowed to update based on prices or are redefined as pure indexers trading the market portfolio.

$LA \cup SA$, and $\mathcal{F}_j = \{\bar{z}\}$ for $j \in LP \cup SP$. In the date 1 asset market, investor j chooses demand $\{q_{ji}\}_{i=1}^n$ for the risky assets to maximize the following mean-variance utility:

$$U_j = E\left[\sum_{i=1}^{n} q_{ji}(\tilde{z}_i - r\tilde{p}_i) \mid \mathcal{F}_j\right] - \frac{\rho}{2} Var\left[\sum_{i=1}^{n} q_{ji}(\tilde{z}_i - r\tilde{p}_i) \mid \mathcal{F}_j\right]. \tag{1}$$

Learning capacity. On date 0, active investors can acquire private signals regarding the risky assets, but the quality of the signals is constrained by each investor's capacity to process information, $K_j \geq 0$, which places a limit on the reduction of uncertainty regarding asset payoffs. $\alpha_{ji} \equiv \frac{Var[\tilde{z}_i]}{Var[\tilde{z}_i|\tilde{s}_{ji}]}$ is defined as an investor j's learning choice for asset i. Following Kacperczyk et al. (2016), we impose a linear capacity constraint such that the sum of the uncertainty reduction must not exceed the information capacity:

$$\sum_{i=1}^{n} \alpha_{ji} \leq n + 2K_j. \tag{2}$$

As evident in (2), a higher capacity K_j implies more resources to gather and process information on different assets, which translates into signals that track realized payoffs with higher precision. For retail and passive investors who lack information capacity ($K_j = 0$ for $j = 0$ or $j \in SP \cup LP$), $\alpha_{ji} = 1$. On date 0, observing other active investors' information choices, active investor j chooses capacity allocation $\{\alpha_{ji}\}_{i=1}^n$ to maximize the ex-ante expected utility $E[U_j]$.

2.2 Equilibrium

The economy is defined by a tuple of exogenous parameters $\mathcal{E} = \{n, l, r, \rho, \{\bar{z}_i\}_{i=1}^n, \{\sigma_i\}_{i=1}^n, \{\bar{x}_i\}_{i=1}^n, \{\sigma_{xi}\}_{i=1}^n, \{K_j\}_{j=1}^l, \{\lambda_j\}_{j=0}^l\}$. An equilibrium consists of active investors' date 0 information-allocation strategies, $\{\alpha_{ji}^*\}_{j=1,\ldots,l;i=1,\ldots,n}$, all investors' date 1 trading strategies $\{q_{ji}(\tilde{s}_{ji}, \tilde{p}_i)\}_{j=1,\ldots,l;i=1,\ldots,n}$, and date 2 price functions $\{\tilde{p}_i\}_{i=1}^n$ such that:

i. active investors' information allocation strategies $\{\alpha_{ji}^*\}_{j=1,\ldots,l;i=1,\ldots,n}$ form a Nash equilibrium: $\alpha_{ji}^* = \arg\max_{\alpha_{ji}} E[\, U_j(q_{ji}(\tilde{s}_{ji}, \tilde{p}_i),\ q_{j'i}(\tilde{s}_{j'i}, \tilde{p}_i))\,]$ where $j, j' \in LA \cup SA$ and $j' \neq j$,

subject to the constraint (2);

ii. the trading strategies $\{q_{ji}(\mathcal{F}_j)\}_{j=1}^{l}$ form a Bayesian-Nash equilibrium in the asset market: $q_{ji}(\mathcal{F}_j) = \arg\max_{q_{ji}} E[\, U_j(q_{ji},\, q_{j'i}(\mathcal{F}_{j'})) \mid \mathcal{F}_j \,]$ for $\forall j$, where $j' \neq j$; and

iii. the price $\tilde{p}_i$ clears the market for asset i, where $i \in \{1, ..., n\}$: $\sum_{j=0}^{l} \lambda_j q_{ji} = \tilde{x}_i$.

Trading decisions. As is standard in the literature (e.g., Kyle, 1989), we consider the following linear demand schedule of investor j for asset i in the date 1 financial market:

$$q_{ji} = \beta_{0ji} + \beta_{1ji}\, \tilde{s}_{ji} - \beta_{2ji}\, r\, \tilde{p}_i, \tag{3}$$

where the β-coefficients are endogenously determined in equilibrium. On date 1, investors choose their demand $\{q_{ji}\}$ to maximize the utility in equation (1). The solution to the problem depends on whether the investor is oligopolistic or retail, with their optimal demand respectively given as follows:

$$q_{ji} = \frac{\mu_{ji} - r\tilde{p}_i}{\rho\hat{\sigma}^2_{ji} + r\frac{d\tilde{p}_i}{dq_{ji}}} \text{for oligopolists and } q_{0i} = \frac{\mu_{0i} - r\tilde{p}_i}{\rho\hat{\sigma}^2_{0i}} \text{ for retail investors,} \tag{4}$$

where the conditional moments $\mu_{ji} \equiv E[\tilde{z}_i|\mathcal{F}_j]$ and $\hat{\sigma}^2_{ji} \equiv Var[\tilde{z}_i|\mathcal{F}_j]$ are computed using Bayes' rule. Here, oligopolists internalize their price impact when making trading decisions, thus introducing a wedge $\frac{d\tilde{p}_i}{dq_{ji}}$ into the otherwise standard CARA demand function. Given the linear demand (3) and the market-clearing condition, the price impact of oligopolist j on asset i is $\frac{d\tilde{p}_i}{dq_{ji}} = \frac{\lambda_j}{r\sum_{k=-j}\lambda_k\beta_{2ki}}$. The demand of retail investors does not move the price and thus takes the standard form.

Given (4) and the market-clearing condition, investors' demand schedules (3) can be summarized by a fixed point of the system, conditional upon active investors' information choices on

date 0:

$$\beta_{0ji} = \frac{-\frac{\gamma_{ji}}{\Delta_i}\left(-\bar{x}_i + \sum_{k=0}^{l}\lambda_k\beta_{0ki} + \sum_{k\neq j}\lambda_k\beta_{1ki}\frac{1}{\alpha_{ki}}\bar{z}\right)}{\rho\hat{\sigma}^2_{ji} + r\frac{d\tilde{p}_i}{dq_{ji}}},$$
$$\beta_{1ji} = \frac{1 - \frac{\gamma_{ji}}{\Delta_i}\left(\lambda_j\beta_{1ji} + \sum_{k\neq j}\lambda_k\beta_{1ki}\left(1 - \frac{1}{\alpha_{ki}}\right)\right)}{\rho\hat{\sigma}^2_{ji} + r\frac{d\tilde{p}_i}{dq_{ji}}},$$
$$\beta_{2ji} = \frac{1 - \frac{\gamma_{ji}}{r}}{\rho\hat{\sigma}^2_{ji} + r\frac{d\tilde{p}_i}{dq_{ji}}},$$

where $\gamma_{ji} \equiv \frac{Cov_j(\tilde{z}_i,\tilde{p}_i)}{Var_j(\tilde{p}_i)}$ is used by investors to update their beliefs after observing prices and $\Delta_i \equiv r\sum_{j=0}^{l}\lambda_j\beta_{2ji}$ is the size-weighted sensitivity of investor demand to prices. For retail investors, there is no price impact, i.e., $\frac{d\tilde{p}_i}{dq_{0i}} = 0$. For passive investors, without any information, $\gamma_{ji} = 0$, and the system simplifies to

$$\beta_{0ji} = 0 \text{ and } \beta_{1ji} = \beta_{2ji} = \frac{1}{\rho\hat{\sigma}^2_{ji} + r\frac{d\tilde{p}_i}{dq_{ji}}}. \tag{5}$$

Information-acquisition decisions. We move backward to date 0 to characterize active investors' information-acquisition decisions (note that $\alpha_{ji} = 1$ for $j \in SP \cup LP \cup \{0\}$). Given other active investors' information choices $\{\alpha_{j'i}\}_{j'\neq j}$, active investor j chooses $\{\alpha_{ji}\}$ to maximize their expected utility $E[U_j]$, subject to constraint (2). In other words, each active investor's information choice is a best response to choices made by other investors. The equilibrium information choices arise from the interplay of these best responses.

Informativeness and concentration. Our analysis centers on the relationship between ownership concentration and price informativeness, i.e., a measure of the amount of information incorporated into asset prices. Following Bai et al. (2016) and Kacperczyk et al. (2025), we measure price informativeness as the covariance of the price with the asset fundamental, normalized by

the volatility of the price:

$$PI_i \equiv \frac{Cov(\tilde{p}_i, \tilde{z}_i)}{\sqrt{Var(\tilde{p}_i)}} = \frac{\sigma_i \sum_{j=0}^{l} \omega_{ji}(1 - \frac{1}{\alpha_{ji}})}{\sqrt{\frac{\sigma_{x_i}^2}{\sigma_i^2} + \left(\sum_{j=0}^{l} \omega_{ji}(1 - \frac{1}{\alpha_{ji}})\right)^2 + \sum_{j=0}^{l} \omega_{ji}^2 \frac{\alpha_{ji}-1}{\alpha_{ji}^2}}}, \tag{6}$$

where $\omega_{ji} \equiv \frac{\partial \lambda_j q_{ji}}{\partial s_{ji}} = \lambda_j \beta_{1ji}$, capturing investors' trading aggressiveness based on their signals. This measure of price informativeness maps well to our framework since the square root of the reduction in the variance of a Bayesian agent's posterior beliefs captures what they learned from the price. In addition, Bai et al. (2016) show that price informativeness can be derived as a welfare measure under the Q-theory.

Based on (6), price informativeness is shaped by two key variables: (i) ω_{ji} captures how an oligopolist's total demand for asset i responds to their private signal s_{ji}, which is referred to as the information pass-through effect by Kacperczyk et al. (2025); (ii) α_{ji} captures an oligopolist's learning choices, which is termed the learning effect by Kacperczyk et al. (2025).

Depending on the category of institutional investors, we can define ownership concentration among active, passive, or total institutional investors. As will be discussed in Section 2.3.3, it is the learning and trading of active investors that drive price informativeness; thus, we will focus on active investors. We explore two layers of ownership concentration. First, we follow Kacperczyk et al. (2025) to measure concentration among active investors based on their size:

$$ActHHI_{size} = \sum_{j \in SA \cup LA} \left(\frac{\lambda_j}{\sum_{k \in SA \cup LA} \lambda_k} \right)^2. \tag{7}$$

Second, we utilize a new ownership concentration measure constructed based on investors' endogenous asset demand $\{q_{ji}\}$:

$$ActHHI_{trading} = \sum_{j \in SA \cup LA} \left(\frac{\lambda_j E[|q_{ji}|]}{\sum_{k \in SA \cup LA} \lambda_k E[|q_{ki}|]} \right)^2. \tag{8}$$

Given the static nature of our model, investors' asset demand can also serve as a proxy for their

asset holdings. Consequently, this measure enables us to leverage extensive firm-level holding data and the resulting variation employed in our empirical analysis.

2.3 Numerical Results

Following Kacperczyk et al. (2025), we conduct numerical analyses to derive the model implications of institutional ownership concentration for price informativeness. These implications are tested in subsequent sections.

2.3.1 Parameter Choices

The values of parameters used in our numerical analysis are summarized in Table 1. Specifically, per Kacperczyk et al. (2025), we set the asset payoff distribution to $\bar{z}_i = 10$ and $\sigma_i = 1$ for all i, the number of assets to $n = 5$, and the number of oligopolists to $l = 20$. Moreover, the volatility of asset supply, σ_{xi}, is chosen with a target coefficient of variation of 0.2 for all i. The risk-free rate is set to match the real return of 2.5% on three-month T-bills. The risk aversion coefficient ρ is 2.32, and the learning capacity is $K_j = 12.5$ for large active oligopolists and $K_j = 1.25$ for small oligopolists. The supply of risky assets, $\bar{x}_i$, is linearly distributed between 3 and 6, featuring a narrower range between the largest and smallest assets compared to the range employed by Kacperczyk et al. (2025). This setup allows for the smallest asset to be learned.

In addition, our investor mass $\{\lambda_j\}_{j=0}^{l}$ is set to match the empirical ownership distribution. Specifically, we chose the retail ownership $\lambda_0 = 40\%$ to reflect the fact that institutional ownership has fluctuated between 55% and 65% over the past two decades, per Thomson Reuters 13F holding data (our results remain robust even when this value is varied between 35% and 45%). The remaining 60% of institutional holdings are allocated among 20 oligopolists.

As in Kacperczyk et al. (2025), half of the oligopolists are active, and the rest are passive. Within the active or passive group, two oligopolists are assumed to be large and the remaining eight are assumed to be small; that is, $LA = \{1, 2\}$, $LP = \{3, 4\}$, $SA = \{5, \ldots, 12\}$, and $SP = \{13, \ldots, 20\}$. Furthermore, the relative size within each small group is set to be linearly

distributed between 1 and 5, meaning the largest small active oligopolist is five times larger than the smallest one. The same is true for small passive oligopolists.

For passive ownership, Kacperczyk et al. (2025) assume that the size of the passive sector is 20% of total institutional ownership, based on the index fund share published in the Investment Company Institute (ICI) Fact Book. However, index funds are not the only type of passive investor. Examining the closing volumes of index additions and deletions on reconstitution days, Chinco and Sammon (2024) estimate that passive investors have held approximately 30% of the U.S. stock market in the past decade. Thus, in our model, with 40% retail ownership, the passive sector is approximately 50% of total institutional ownership, i.e., $\sum_{j\in SP\cup LP}\lambda_j/\sum_{j=1}^{l}\lambda_j = 30\%/60\% = 50\%$ (our results remain robust even when the value is varied between 20% and 60%).

Finally, to study the effect of institutional ownership concentration on price informativeness, we follow Kacperczyk et al. (2025) and generate different concentration levels by varying two sets of parameters to produce an increasing HHI index for oligopolistic ownership. First, we vary the relative size of the two large active oligopolists and the two large passive oligopolists by increasing the ratios λ_1/λ_2 and λ_3/λ_4 linearly from 1.1 to 10 across ten scenarios. Second, and simultaneously, we vary the relative size of small sectors, given by $\sum_{j\in SA}\lambda_j/\sum_{j\in SA\cup LA}\lambda_j$ and $\sum_{j\in SP}\lambda_j/\sum_{j\in SP\cup LP}\lambda_j$, linearly from 10% to 3% across the same ten scenarios. These parameter changes mirror the empirical trend of an increasing relative size within the large sector and, concurrently, a decreasing relative size within the small sector over time. The resulting distribution of λ_j across the ten calibration scenarios closely matches the data. Section A of the online appendix provides further details.

2.3.2 Ownership Concentration and Price Informativeness

To begin, Figure 2 plots how size-based ownership concentration $ActHHI_{size}$ among active oligopolists, as defined in equation (7), affects price informativeness on an asset-by-asset basis. Consistent with Figure 10 of Kacperczyk et al. (2025), Panel (a) of Figure 2 shows that price informativeness of all assets decreases with higher size-based active ownership concentration.

The following hypothesis summarizes this result.

Hypothesis 1 (Size-based concentration)**.** *Price informativeness is lower when size-based active ownership concentration is higher.*

As large active investors increase in size, they tend to trade conservatively based on their private information; meanwhile, smaller active investors exhibit the opposite behavior as they decrease in size. This shift in size distribution also influences how active investors allocate their capacity for information learning. Together, these factors shape price informativeness through the investors' trading and learning decisions. Section 2.3.4 offers a detailed discussion of the underlying mechanisms.

We next explore trading-based ownership concentration $ActHHI_{trading}$, as defined in equation (8), in Figure 3. This measure is specific to each asset. Analyzing all panels of Figure 3 for assets of varying sizes, we observe that a higher concentration of trading-based ownership is associated with a decline in individual price informativeness. Given the endogenous nature of this measure, which is calculated based on trading volume as shown in equation (8), the underlying driver of its effect on price informativeness must stem from the investor size λs.

To illustrate this point, Figure 4 plots the trading activity of the largest active investor alongside that of other investors. Recall that an increase in $ActHHI_{size}$ is driven by an increase in the size λ_1 of the largest active institutional investor and a decrease in the sizes of other active institutional investors (see Table A1). Panel (a) shows that an increase in the size of the largest active investor leads said investor to heighten their trading activity across all risky assets, resulting in varying increases in the trading volume of different assets. Conversely, when the sizes of other active investors decrease, the trading volume declines across all assets, as illustrated in Panel (b). So, as $ActHHI_{size}$ increases, the largest active investor is responsible for more trades per asset, which implies that trading-based concentration $ActHHI_{trading}$ increases with size-based concentration $ActHHI_{size}$, as shown in Panel (c). Combined with size-based ownership concentration's negative effect on price informativeness, our findings, as summarized by the following hypothesis, suggest that trading-based ownership concentration is also negatively related to individual

price informativeness.

Hypothesis 2 (Trading-based concentration)**.** *Price informativeness is lower when trading-based active ownership concentration is higher.*

2.3.3 Active Versus Passive Ownership Concentration

Our analysis thus far has demonstrated that, as the concentration of active ownership rises, price informativeness diminishes. However, since we follow Kacperczyk et al. (2025) and increase the ownership concentration of both active and passive investors in our numerical analysis, there is a concern that the reduced price informativeness might be attributed to increasing ownership concentration among passive rather than active investors.

Although passive investors lack learning capacity, they can indirectly influence price informativeness by affecting the trading decisions of active investors. For example, as large passive investors increase in size while small passive investors diminish in size, the former group becomes less responsive to price changes, whereas the latter group exhibits greater responsiveness (as shown in equation (5), the coefficient β_2 decreases as price impact $d\tilde{p}_i/dq_{ji}$ increases). Due to the increased size of large passive investors, the overall passive sector tends to trade less aggressively in response to price changes. Consequently, active investors become increasingly concerned about their own price impact and thus trade less aggressively based on their private information, which ultimately leads to a decrease in price informativeness.

To evaluate the effect of passive ownership concentration, we conduct two placebo tests by respectively fixing the size distribution of active and passive investors throughout the experiment. For the first placebo test (and a similar approach for the second), the relative size of the two large active investors, λ_1/λ_2, is fixed at 1.1, as in the first scenario. Meanwhile, the relative size of the small active sector, $\sum_{j\in SA}\lambda_j/\sum_{j\in SA\cup LA}\lambda_j$, is fixed at 0.10. This ensures that the ownership concentration among active investors remains constant across ten scenarios, while the ownership concentration among passive investors increases.[3]

[3]Passive ownership concentration is defined similarly to equations (7) and (8) but with active investors replaced

Figures 5 and 6 illustrate the results of our placebo tests using size- and trading-based ownership concentration, respectively. Regardless of the measure, as passive ownership concentration increases, price informativeness remains relatively unchanged. In contrast, when active ownership concentration increases, price informativeness decreases, following a pattern quantitatively similar to what is observed in Figures 2 and 3. This indicates that it is indeed *active* ownership concentration that drives the observed results.

2.3.4 Mechanisms

As do Kacperczyk et al. (2025), we attribute the effect of ownership concentration on price informativeness to two effects pertaining to investors' trading and learning behaviors: the information pass-through effect, which refers to investors' trading aggressiveness against their private signals, represented by $\omega_{ji} \equiv \frac{\partial \lambda_j q_{ji}}{\partial s_{ji}} = \lambda_j \beta_{1ji}$; and the learning effect, which concerns investors' information acquisition choices, captured by parameter α_{ji}. We examine the information pass-through effect by fixing the degree of learning (α_{ji}) at the level defined in the first scenario of the concentration experiment, and we examine the learning effect by fixing the information pass-through (ω_{ji}) at values from the same first scenario. Panels (a) and (b) of Figure 7 present these results.

Information pass-through effect. As shown in Panel (a), when learning (α_{ji}) is fixed, as large active investors grow in size, they trade more conservatively against their private signals (captured by β_1) due to increased concerns regarding the price impact. Meanwhile, small active investors may trade more aggressively against their private signals as their price impact concerns decrease, although their economic importance diminishes as they shrink in size (captured by λ). Taken together, the dropping information pass-through drives the price informativeness down. These results are summarized as follows:

Hypothesis 3 (Information pass-through effect)**.** *When active ownership concentration is higher,*

by passive investors. Formally, $PasHHI_{size} = \sum_{j \in SP \cup LP} \left(\frac{\lambda_j}{\sum_{k \in SP \cup LP} \lambda_k} \right)^2$.

large active investors tend to trade more conservatively based on their information, leading to lowered average trading intensity across all active investors.

Learning effect. When information pass-through (ω_{ji}) is fixed, large active investors diversify their learning as they grow in size. In contrast, small active investors, decreasing in size, specialize their learning in large assets. The pattern in Panel (b) of Figure 7 suggests that this specialized learning dominates.

To clarify why small active investors' specialized learning dominates under the learning effect, Figure 8 compares the learning choices of the largest active investors with those of other active investors. As the largest active investor grows, they stretch their learning capacity across various assets, which increases the price informativeness of smaller assets (assets 1 and 2) and decreases that of larger assets (assets 4 and 5). Conversely, other active investors shrink in size and thus focus their learning capacity on larger assets (assets 4 and 5), which reduces the price informativeness of smaller assets (assets 1 and 2) and increases that of said larger assets (assets 4 and 5). Since the largest active investor has already diversified their learning in the beginning, further diversification has a marginal impact on price informativeness. That is, small active investors' specialized learning prevails, leading to increased price informativeness for large assets and decreased price informativeness for small assets, as observed in Panel (b) of Figure 7. These results are summarized as follows:

Hypothesis 4 (Learning effect)**.** *Higher active ownership concentration leads, on average, to increased learning in large assets and diminished learning in small assets.*

2.3.5 Real Investment

Having established that price informativeness decreases as active ownership concentration increases, a natural follow-up question is whether this reduced informativeness extends to real firm decisions, as discussed by Bond et al. (2012) and Goldstein (2023). Due to the complexity of our framework, we do not explicitly model how asset prices influence managers' investment

decisions, but we adopt Subrahmanyam and Titman (2001) approach and assume that, as prices become less informative, managers predict investment decisions less accurately. Based on this assumption, along with Hypotheses 1 and 2, we anticipate that higher ownership concentration, whether based on size or trading, is likely associated with a lower association between prices and future investments. This leads us to the following hypothesis:

Hypothesis 5 (Real investment)**.** *When active ownership concentration is higher, prices become less informative regarding future investments.*

3 Data and Measures

3.1 Sample Construction

Our main sample includes U.S.-listed companies with common stocks traded on the NYSE, NASDAQ, and AMEX. Firm-level financial statement data is sourced from Compustat, supplemented with the intangible capital estimates as defined in Peters and Taylor (2017) from WRDS, and stock price information is obtained from CRSP.

To construct measures of ownership concentration, we first extract institutional holdings information from the Thomson Reuters 13F database and merge the holdings data with Bushee's (1998) classification scheme to identify active institutional investors. We then construct ownership-related variables, such as active/passive institutional ownership concentration, for each firm or market quarter.

For empirical tests involving size-based ownership concentration, we construct ownership-related variables in the fourth quarter and merge them with price-informativeness measures, which we introduce shortly. For empirical tests involving trading-based ownership concentration, we construct ownership-related variables most recent to the end of each firm's fiscal year and merge them with the firm-level financial statement data. Our samples are all of annual frequency and cover the period from 1980 to 2022.

Our sampling criteria are as follows. Observations with a stock price below 1 dollar, as well as

observations with a market capitalization below 5 million, are excluded, as are firms within the financial industry and those with less than four successive years of accounting data. Furthermore, sample firms are required to have at least one active institutional investor. For empirical tests using trading-based concentration metrics, the requirement is further tightened to sample firms with at least five active institutional investors to avoid extreme values of concentration, though our results persist even when there is only one active institutional investor. Unless otherwise stated, our sample selection criterion remains consistent throughout all subsequent empirical analyses.

Table 2 reports the summary statistics of the variables used in our main analysis. Table B1 of the online appendix provides a comprehensive list of variable definitions. To mitigate the influence of outliers, all continuous variables are winsorized at the 1% and 99% levels.

3.2 Measures of Active Institutional Ownership Concentration

We follow Bushee (1998) to categorize institutions as active or passive investors based on their historical investment behaviors.[4] Specifically, there are three categories of investors: (i) quasi-indexers with low turnover and high diversification, (ii) transient investors with high turnover and high diversification, and (iii) dedicated investors with low turnover and low diversification. As in Kacperczyk et al. (2021), we classify transient and dedicated investors as active and quasi-indexers as passive.

We first construct the size-based ownership concentration measure. In each quarter, we calculate the AUM of each active institutional investor by adding up their holding value in their underlying securities. The first concentration measure refers to the HHI of AUM among active

[4]Bushee's classification has two versions, "permanent" and "time-varying." Following Appel et al. (2016), we use the permanent version in our baseline results to prevent an institutional investor from being classified as active at some points but passive at others. As shown in Table C3 of the online appendix, our results hold even if we use the time-varying classification. Moreover, "institution" here is at the management company level as recorded in 13F holdings data. This poses a challenge since a fund-management company may oversee both passive and active mutual funds, potentially leading to measurement errors in the classification. To address this issue, we utilize fund-level holdings data from Thomson Reuters S12 as an alternative source to distinguish between active and passive mutual funds. Table C2 of the online appendix shows that the results are robust.

institutional investors:

$$ActHHI_{size,q} = \frac{\sum_{j=1}^{N}\left(AUM_{j,q}^{2}\right)}{\left(\sum_{j=1}^{N} AUM_{j,q}\right)^{2}}, \tag{9}$$

where $AUM_{j,q}$ is active institutional investor j's AUM in quarter q, and N is the total number of active institutional investors. The second concentration measure calculates the proportion of AUM held by the top five active institutional investors relative to the total AUM of all active institutional investors:

$$ActTop5_{size,q} = \frac{\sum_{j=1}^{\text{Top 5}} AUM_{j,q}}{\sum_{j=1}^{N} AUM_{j,q}}. \tag{10}$$

We then construct the trading-based ownership concentration measure similar to the size-based measure but using firm-level holding data:

$$ActHHI_{i,q} = \frac{\sum_{j=1}^{N_i}\left(S_{i,j,q}^{2}\right)}{(\sum_{j=1}^{N_i} S_{i,j,q})^{2}} \text{ and } ActTop5_{i,q} = \frac{\sum_{j=1}^{\text{Top 5}} S_{i,j,q}}{\sum_{j=1}^{N_i} S_{i,j,q}}, \tag{11}$$

where $S_{i,j,q}$ denotes the equity shares of stock i owned by active institution j in quarter q, and N_i is the number of active institutions holding stock i.

Both $ActHHI$ and $ActTop5$, whether based on size or trading, are designed to have values between 0 and 1, with 0 representing highly dispersed ownership and 1 representing highly concentrated ownership.

3.3 Measures of Price Informativeness

Adopted from Bai, Philippon, and Savov (2016), our primary measure of price informativeness is welfare-based and maps well to our theoretical framework. In Section C.1 of the online appendix, we explore several alternative measures of price informativeness, including post-earnings-announcement drift (Ball and Brown, 1968), conditional probability of an information event (Duarte et al., 2020), informed trading intensity (Bogousslavsky et al., 2024), variance ratio (Lo and MacKinlay, 1988), and relative price informativeness (Dávila and Parlatore, 2025). Our results, as shown, are robust.

Following Bai et al. (2016), we first run cross-sectional regressions of future earnings on cur-

rent market prices for each year:

$$\frac{E_{i,t+h}}{A_{i,t}} = a_{t,h} + b_{t,h} \log\left(\frac{M_{i,t}}{A_{i,t}}\right) + c_{t,h}\left(\frac{E_{i,t}}{A_{i,t}}\right) + d^{s}_{t,h}\mathbf{1}^{s}_{i,t} + \epsilon_{i,t,h}, \tag{12}$$

where h denotes the prediction horizons, which equals 1 or 3 in our study; $\mathbf{1}^{s}_{i,t}$ denotes a sector indicator defined as the one-digit SIC code; $M_{i,t}/A_{i,t}$ denotes the market price of firm i in fiscal year t, computed as the market capitalization at the end of March after year t, scaled by total assets in year t; and $E_{i,t+h}/A_{i,t}$ ($E_{i,t}/A_{i,t}$) denotes future (current) earnings, computed as cash flow in year $t+h$ (t) scaled by total assets in year t. Following Bai et al. (2016), we measure firm cash flows by using earnings before interest and taxes ($EBIT$); earnings before interest, taxes, depreciation, and amortization ($EBITDA$); and net income (NI). The average price informativeness in year t against prediction horizon h is then calculated as the forecasting coefficient $b_{t,h}$ in equation (12) multiplied by $\sigma_t(\log(M/A))$, the cross-sectional standard deviation of the scaled market price $\log(M/A)$ in year t:

$$PI_{t,h} = b_{t,h} \times \sigma_t(\log(M/A)). \tag{13}$$

Similarly, we estimate investment-to-price sensitivity by first running cross-sectional regressions of future investment rates on current market prices for each year and then multiplying the forecasting coefficient by $\sigma_t(\log(M/A))$:

$$\frac{I_{i,t+h}}{K_{i,t}} = a_{t,h} + b_{t,h} \log\left(\frac{M_{i,t}}{A_{i,t}}\right) + c_{t,h}\left(\frac{E_{i,t}}{A_{i,t}}\right) + d_{t,h}\left(\frac{I_{i,t}}{K_{i,t}}\right) + e^{s}_{t,h}\mathbf{1}^{s}_{i,t} + \epsilon_{i,t,h}, \tag{14}$$

where $I_{i,t+h}/K_{i,t}$ denotes investment rates as in Peters and Taylor (2017), including the intangible investment rate ($Intangible/K$), physical investment rate ($Physical/K$), and total investment rate ($Invest/K$). Specifically, the intangible investment rate ($Intangible/K$) is calculated as $R\&D + 0.3 \times SG\&A$ expenses,[5] scaled by total capital (K), where total capital is defined as the

[5]Only a small proportion of SG&A is related to investments in intangible organization capital. The rest of SG&A relates to operating costs that support the current period's profits. The 30% is a rule of thumb used in prior studies

sum of net property, plant, and equipment (item PPENT from Compustat) and intangible capital (item K_INT from Peters and Taylor (2017)). The physical investment rate ($Physical/K$) is calculated by dividing capital expenditures (CAPX) by total capital. Finally, the total investment rate ($Invest/K$) is the aggregate of the intangible and physical investment rates. The predicted variation of real investments based on prices in year t against prediction horizon h is then calculated as the forecasting coefficient $b_{t,h}$ in equation (14) multiplied by $\sigma_t(\log(M/A))$.

By conducting the cross-sectional regressions for each year, we can estimate a time-series set of price informativeness and investment-to-price sensitivity and examine their relationship using size-based active ownership concentration. However, the cross-sectional nature of this estimation makes it unsuitable to utilize investors' firm-level equity holding data, given that the holding data is panel data while price informativeness is time-series data. Kacperczyk et al. (2021) address this issue by modifying the cross-sectional regression into a pooled regression. We thus use the cross-sectional regression model to estimate price informativeness when studying the effect of size-based concentration and the modified pooled regression model detailed in Section 5.1 when studying the effect of trading-based concentration.

4 Size-based Ownership Concentration

In this section, we follow an approach developed by Bai et al. (2016) and Farboodi et al. (2022) to test Hypotheses 1 and 5. Specifically, we investigate the effect of size-based active ownership concentration on price informativeness and how current prices predict future investments. This analysis directly tests the implications suggested by Kacperczyk et al. (2025).

First, we visually inspect the relationship between size-based active ownership concentration and price informativeness estimated from equation (13) using a one-year prediction horizon.[6] Figure 9 presents scatter plots along with the fitted lines and 95% confidence intervals. Panels (a1) through (a3) use $ActHHI_{size}$ in equation (9) to measure concentration while Panels (b1) through

(Eisfeldt and Papanikolaou, 2014; Peters and Taylor, 2017; Jha et al., 2024).

[6]This negative relation remains robust when we change the prediction horizon from one year to three years, as shown in Figures B1 and B2 in the online appendix.

(b3) use $ActTop5_{size}$ in equation (10) as an alternative measure. A significantly negative correlation is observed between size-based active ownership concentration and price informativeness across different specifications, consistent with Hypothesis 1 and its numerical analysis in Figure 2. Moreover, the effect is economically meaningful. For example, the correlation coefficient is -0.18 in Panel (a2) of Figure 9, suggesting that a 1%-point increase in $ActHHI_{size}$ is associated with a 25.7% decrease in price informativeness relative to its mean level of 0.007.

Second, we divide the sample firms into five groups based on their market capitalization and estimate each group's price informativeness. Figure 10 presents scatter plots along with fitted lines and reveals two noteworthy observations: First, larger firms enjoy higher price informativeness on average, consistent with our numerical results in Figure 2. This is also consistent with Farboodi et al. (2022), who show that data-processing efforts in large firms are much higher than those in small firms. Second, the negative correlation between size-based ownership concentration and price informativeness holds for all size groups, suggesting that our results are not driven by any specific group of firms.

Next, we examine how size-based active ownership concentration affects investment-to-price sensitivity. Consistent with Hypothesis 5, a negative correlation is observed across all specifications in Figure 11, which is generally statistically significant except for Panel (b1). This implies that, as active ownership concentration suppresses price informativeness, the ability for prices to predict future investments also diminishes.

Despite the significance and robustness of our results, we remain cognizant of their limitations. For instance, the sample size is small due to the low data frequency (42 for $h = 1$ and 40 for $h = 3$), indicating that the point estimates may be sensitive to different empirical setups. Furthermore, some estimates of price informativeness in Figures 9 and 10 are negative, which previous studies using similar estimation processes have also observed, e.g., Dávila and Parlatore (2025) and Farboodi et al. (2022). These limitations thereby motivate and justify further intensive explorations at the firm level.

5 Trading-based Ownership Concentration

In this section, we use a method developed by Kacperczyk et al. (2021) to investigate the effect of our newly proposed trading-based active ownership concentration. Specifically, Section 5.1 conducts baseline regressions, Section 5.2 addresses the issue of endogeneity by examining the context of mergers among active financial institutions, and Section 5.3 expands our sample to an international setting. Section C of the online appendix provides additional robustness analyses.

5.1 Baseline Regressions

To examine the effect of trading-based active ownership concentration on price informativeness as stated in Hypothesis 2, we follow Kacperczyk et al. (2021) and estimate the following pooled regression model using firm-level data at an annual frequency:

$$\begin{aligned}\frac{E_{i,t+h}}{A_{i,t}} = a_h + b_h \log\left(\frac{M_{i,t}}{A_{i,t}}\right) + c_h \log\left(\frac{M_{i,t}}{A_{i,t}}\right) \times Concentration_{i,t} + d_h Concentration_{i,t} \\ + e_h \frac{E_{i,t}}{A_{i,t}} + f_h \chi_{i,t} + g_h \log\left(\frac{M_{i,t}}{A_{i,t}}\right) \times \chi_{i,t} + FE_{i,t} + \varepsilon_{i,t+h},\end{aligned} \tag{15}$$

where h denotes the prediction horizons, equaling 1 or 3; $Concentration_{i,t}$ denotes the trading-based ownership concentration among active institutional investors, measured by $ActHHI$ or $ActTop5$, as defined in equation (11); and $E_{i,t}/A_{i,t}$ is one of the three measures of earnings ($EBIT$, $EBITDA$, and NI), scaled by total assets. $\chi_{i,t}$ is a saturated set of control variables: passive ownership concentration ($PasHHI$ or $PasTop5$), calculated in the same way as active ownership concentration, except that we use the holding information of passive institutional investors; institutional ownership (IO), calculated as the total share holdings by institutional investors divided by market capitalization; firm leverage ($Leverage$), defined as book debt divided by total assets; firms' total sales scaled by total assets ($Sale$); firms' cash holdings scaled by total assets ($Cash$). We include firm fixed effects to control for unobserved, omitted firm characteristics correlated with both ownership concentration and price informativeness measures. We also include industry-year fixed effects to absorb time-varying economic or regulatory shocks at the

industry level (Antón et al., 2023), defined based on the first two digits of SIC codes. $\varepsilon_{i,t+h}$ is the error term, double clustered at both the firm and year levels to account for possible dependence along those two dimensions. The coefficient c_h is of interest, measuring average price informativeness defined as the sensitivity of future earnings to current stock prices, conditional upon the active institutional ownership concentration.

Panel A of Table 3 uses $ActHHI$ to measure active institutional ownership concentration. In Columns (1) through (3), we use scaled $EBIT$, $EBITDA$, and NI to measure earnings, respectively. The coefficient of interest, $c_{h=1}$, is statistically significantly negative at the 1% level. The effect is also economically significant: for example, $c_{h=1} = -0.030$ in Column (2), indicating that, when $ActHHI$ increases from the 25th to the 75th quantiles while other control variables stay constant at their mean levels, price informativeness decreases by 24.2%. In Columns (4) through (6), we perform the same estimation regression for price informativeness but against a three-year prediction horizon. The coefficients $c_{h=3}$ remain significantly negative and somewhat larger in magnitude: for example, the coefficient $c_{h=3} = -0.059$ implies that, when $ActHHI$ increases from the 25th to the 75th quantile, conditional on other control variables staying constant at their mean levels, price informativeness decreases by 40.6%.

Panel B of Table 3 replicates the results in Panel A but employs $ActTop5$ to measure active institutional ownership concentration. A significantly negative effect on price informativeness continues to be observed, with the economic magnitude comparable to that in Panel A: for instance, the coefficients $c_{h=1} = -0.040$ and $c_{h=3} = -0.063$ suggest that an interquartile range move in $ActHHI$, with other control variables held constant at their mean levels, corresponds to a decrease of 27.8% and 45.2% in price informativeness against a one- and three-year prediction horizon, respectively. These results are consistent with Hypothesis 2 in Section 2.3.

We then estimate the effect of active institutional ownership concentration on investment-to-price sensitivity in a similar fashion to regression (15) but with the scaled earnings E/A replaced by investment rate I/K. Panel A of Table 4 uses $ActHHI$ to measure active institutional ownership concentration while Panel B uses $ActTop5$. The coefficients on the interaction term,

$\log(M/A) \times Concentration$, are negative and statistically and economically significant across different specifications. Take the results related to physical investment in Columns (2) and (5) of Panel A in 4, for example: the coefficients $c_{h=1} = -0.023$ and $c_{h=3} = -0.025$ suggest that when $ActHHI$ rises from the 25th to the 75th quantile, with other control variables held constant at their mean levels, investment-to-price sensitivity decreases by 10.7% and 12.6% against a one- and three-year predication horizon, respectively. These results suggest that, consistent with Hypothesis 5 in Section 2.3, the predictive power of the current stock price for future investment decisions is weakened for firms with more concentrated active institutional ownership.

Finally, controlling for passive ownership concentration under all specifications did not alter the negative effect of active ownership concentration on price informativeness. In general, the effect of passive ownership concentration on price informativeness is empirically inconsistent across specifications: for example, using post-earnings-announcement drift as an alternative measure of price informativeness (Table C1 in the online appendix), we find that passive ownership concentration is associated with increased price informativeness, contrary to the results presented in Table 3. Similarly, when utilizing S12 mutual fund holdings to construct the concentration measure (Table C2 in the online appendix), the coefficients on the interaction terms $\log(M/A) \times PasHHI$ and $\log(M/A) \times PasTop5$ are unstable, fluctuating between positive and negative and varying from significant to insignificant. These inconsistent results largely align with our model, indicating that passive ownership concentration does not exert a meaningful impact on price informativeness.

5.2 Identification

There is a potential concern that the observed negative effect of active ownership concentration may be attributable to unobservable economic forces correlated with both ownership concentration and price informativeness. Another concern is that firms with lower price informativeness, and consequently greater exploitable mispricing opportunities, may attract more institutional blockholders. We address these potential endogeneity issues utilizing a quasi-natural experiment

capturing financial institution mergers, thus generating plausibly exogenous variation in a firm's ownership structure. Our DiD estimation methodology not only attenuates the endogeneity issue but also addresses concerns of measurement error, given that the estimation of $b_{3,h}$ does not rely on ownership concentration measures.

As He and Huang (2017) elucidated, an experiment with institutional mergers hinges on the premise that the reasons for mergers are often unrelated to the fundamentals of portfolio holdings. Upon merging, the acquirer typically assumes control of the target's existing portfolios and, owing to liquidity and transaction cost considerations, retains these acquired holdings for an extended duration. Consequently, if a firm is held by both an active acquirer and an active target prior to the merger, we anticipate an exogenous surge in its active institutional ownership concentration immediately following the merger.

We assemble a sample of financial institution mergers, adhering broadly to the criteria delineated in the literature on cross-ownership (e.g., He and Huang, 2017; Lewellen and Lowry, 2021; Levonyan and Mengano, 2024). First, we retrieve all mergers announced between 1980 and 2021 from the SDC mergers and acquisitions database. Second, we require that (i) the target firm be incorporated in the U.S., (ii) both the acquirer and target be in the finance industry, and (iii) firm names be accessible for both merger participants. Third, for each target and acquirer firm across these deals, we employ text-matching algorithms to align firm names with 13F data.[7] Upon merging the SDC and 13F data, we further require that either the target firm ceased filing 13F statements within 15 months of the merger's completion date or the target's AUM diminished by over 80% from quarter -6 to quarter 6 relative to the completion quarter.

In addition to these data-cleaning procedures, several modifications are implemented to align the setting more closely to our research focus. Specifically, the acquirer's AUM is required to exceed 100 million dollars and increase by at least 1.5 times from quarter -6 to quarter 6 relative to the completion quarter. Also, per Bushee's classification, both merger partners are required to

[7]SDC provides firm names in three forms: the Company, the Immediate Parent Company, and the Ultimate Parent Company. While the names are largely identical for most companies, discrepancies can arise. We utilize all three names in matching SDC mergers with 13F data.

be active. This process yields a sample of 11 active financial institution mergers, as detailed in Table B2 of the online appendix.

For each of the 11 mergers, treated firms are defined as those held by both partners prior to the merger announcement.[8] To preclude trivial holding positions, each partner's holding value is required to exceed 0.01% of the stock's market capitalization. Meanwhile, control firms are defined as those held by either the acquirer or target, amounting to at least 0.01% of the market capitalization before the merger announcement. This strategy for selecting control firms accounts for institutional heterogeneity, such as managerial styles or abilities (Kini et al., 2024).[9] To mitigate potential estimation bias stemming from the "bad comparisons" problem discussed by Baker et al. (2022), firms in the control group that had been treated by any of the other merger events are excluded. The final sample includes 700 unique treated firms and 2,130 unique control firms. To focus on merger shocks, we restrict our analysis to a window of two years before and two years after the merger.[10]

We first check the validity of our DiD design by examining whether active financial institution mergers induce significant increases in active institutional ownership concentration. Specifically, we run the following regression model on a quarterly basis:[11]

$$Concentration_{i,q} = \alpha + \beta Post_q \times Treat_i + Merger \times FE_{i,q} + \epsilon_{i,q},$$

where $Concentration_{i,q}$ denotes the trading-based ownership concentration among active institutional investors, measured by $ActHHI$ and $ActTop5$; $Treat$ is a dummy variable equal to 1 for treated firms and zero for control firms; $Post$ is, for any given merger event, a dummy variable equal to one for the merger completion quarter and all quarters after and zero for the quarters

[8] We utilize the announcement date, rather than the completion date, to ensure that the treatment and control samples are defined using only ex-ante information (Lewellen and Lowry, 2021).

[9] In Table B3 of the online appendix, we consider an alternative strategy for selecting control firms, whereby control firms are defined as those held by the acquirer but not the target, with a 0.01% or greater ownership prior to the merger announcement. The results remain virtually identical.

[10] Table B4 of the online appendix shows that our results remain robust against an alternative estimation window of three years before to three years after the merger.

[11] The results remain robust if we calculate the mean values of $Concentration_{i,q}$ within each firm-year and re-run the regression model on an annual basis.

before; and $Merger \times FE_{i,q}$ denotes the merger-firm and merger-quarter fixed effects, per He and Huang (2017). Our regression model, given the "never-treated" requirement on the control group, aligns with the stacked regression estimator approach discussed by Baker et al. (2022) and adopted in recent studies (e.g., Cengiz et al., 2019; Hollingsworth et al., 2024). Standard errors, $\epsilon_{i,q}$, are clustered two ways at the firm and quarter levels.

Panel A of Table 5 reports the results on post-merger changes in the two concentration measures, $ActHHI$ and $ActTop5$. Both concentration measures significantly increase following active financial institution mergers across different event windows, specifically quarters $(-8, +8)$ and $(-12, +12)$. Thus, active financial institution mergers provide a valid quasi-natural experiment, creating an exogenous and positive shock to ownership concentration among active institutional investors.

Next, we investigate the merger shock's impact on price informativeness by estimating the subsequent regression model against annual accounting information:

$$\begin{aligned} E_{i,t+h}/A_{i,t} = a &+ b_{1,h}\log(M/A)_{i,t} + b_{2,h}Treat_i \times Post_t + b_{3,h}\log(M/A)_{i,t} \times Treat_i \times Post_t \\ &+ b_{4,h}\log(M/A)_{i,t} \times Treat_i + b_{5,h}\log(M/A)_{i,t} \times Post_t \\ &+ b_{6,h}\chi_{i,t} + b_{7,h}\log(M/A)_{i,t} \times \chi_{i,t} + Merger \times FE_{i,t} + \varepsilon_{i,t+h}. \end{aligned}$$

$Post$ is, for any given merger event, a dummy variable equal to one for the merger completion year and all years after and equal to zero for the years before. We include merger-firm and merger-year fixed effects to absorb time-invariant characteristics across firms within the same merger as well as time-varying common time trends across mergers. Standard errors are clustered by firm and year. The regression model used to estimate the shock's effect on investment-to-price sensitivity is similar, except the cash flow variables E/A are replaced with the investment variables I/K. Our coefficient of interest is $b_{3,h}$, which measures this sensitivity around the treatment group's shock relative to that of the control group.

Panels B and C of Table 5 present our results. The price informativeness of the treated firms

diminishes significantly following the shock, along with a reduction in the predictive power of the current stock price for future investment decisions. Additionally, the effect is generally larger against a three-year prediction horizon compared to the one-year short-run horizon.

Figure 12 plots the estimated effect of the merger shock over time in an extended window of $(-3, +3)$ years, with year-0 denoting the merger completion year. Panel (a) measures price informativeness based on the earnings variable $EBITDA/A$ against a one-year prediction horizon while Panel (b) measures investment-to-price sensitivity based on $Invest/K$ under the same horizon. Notably, the negative effect of active financial institution mergers is absent prior to the merger shock, given that the estimated coefficients are indistinguishable from zero prior to the year of merger completion. This observation supports the plausibility of the parallel trend assumption. Furthermore, the negative effect is gradual, increasing in magnitude over time following the merger completion year without exhibiting any reversal. Overall, our DiD estimation results suggest that the negative effect of firms' active institutional ownership concentration is causal.

5.3 International Evidence

In this subsection, we examine whether the negative impact of active institutional ownership concentration on price informativeness prevails in markets outside of the U.S.

We construct an international sample by combining global institutional ownership data from FactSet, accounting data from Worldscope, and stock market data from DataStream. The sample has an annual frequency and spans from 2000 to 2022. We exclude firms within the financial industry and require that a firm possess market capitalization above $1 million and have a minimum of five active institutional investors. We further restrict our sample to countries with at least 20 firms possessing adequate financial information. The final sample includes 22,887 unique firms across 63 countries.

Within the online appendix, Table B5 provides descriptive statistics while Figure B3 displays the time-series average trading-based $ActTop5$ values for the largest equity markets globally.

Note that the average $ActTop5$ value in the U.S. hovers around 50% yet remains the lowest among the nine markets examined. Conversely, markets such as China, Japan, and Australia exhibit higher average $ActTop5$ values at approximately 80% over the last decade. This observation underscores the significance of active institutional ownership concentration on a global scale. Furthermore, $ActTop5$ is notably high at the onset of the sample period, a result that may be attributed to the relatively limited coverage of institutional holdings in FactSet in the early 2000s.

We follow the classification criteria in Kacperczyk et al. (2021) to identify active and passive institutional investors within the international sample, with active investors covering mutual funds, investment advisors, and hedge funds and passive investors covering all remaining types, i.e., bank trusts, insurance companies, pension funds, endowments, index funds, and ETFs. The regression model closely mirrors that of the U.S. sample, with the difference being the incorporation of country-year fixed effects in lieu of industry-year fixed effects. This adjustment aims to absorb country-level economic or regulatory fluctuations across time, though our results remain unchanged when even using industry-year fixed effects.

Table 6 shows that the coefficients on the interaction term are significantly negative in all specifications, suggesting that the negative impact of active institutional ownership concentration on price informativeness, as well as investment-to-price sensitivity, persists in an international setting. To ensure that the observed negative effect is not purely driven by firms located in the U.S., we exclude all U.S. firms and demonstrate the consistent negative impact in Table B6 of the online appendix.

6 Mechanisms

Section 2.3 explores two underlying channels through which active ownership concentration might undermine price informativeness. In this section, we test (i) the "information pass-through effect" (Hypothesis 3), which suggests that heightened price impacts induce larger investors to trade more conservatively on their private information, leading to a decrease in average trading

intensity as active ownership concentration increases; and (ii) "the learning effect" (Hypothesis 4), which posits that increased ownership concentration leads larger investors to diversify their learning while smaller investors specialize. Within the online appendix, we rule out several alternative explanations in Section D and explore competing hypotheses in Section E.

6.1 Information Pass-through Effect

Portfolio turnover. If the information pass-through channel is valid, an active institutional investor would be expected to make smaller position adjustments in stocks if said investor is among the top five largest shareholders, compared to if they only hold a minor stake.

To test this hypothesis, we categorize the holding portfolio of each institutional investor into two subgroups: Top5, which includes stocks for which the investor is one of the top 5 largest shareholders, and Non-Top5, which includes all of the investor's remaining stocks. We then construct the portfolio turnover measures following Yan and Zhang (2009). For each investor k in each quarter q, we first calculate the aggregate purchase and sale for each subgroup g as follows:

$$AgBuy_{k,g,q} = \sum_{i \in N_{k,g}} \left| S_{k,g,i,q} P_{i,q} - S_{k,g,i,q-1} P_{i,q-1} - S_{k,g,i,q-1} \Delta P_{i,q} \right|, \text{ where } S_{k,g,i,q} > S_{k,g,i,q-1},$$

$$AgSell_{k,g,q} = \sum_{i \in N_{k,g}} \left| S_{k,g,i,q} P_{i,q} - S_{k,g,i,q-1} P_{i,q-1} - S_{k,g,i,q-1} \Delta P_{i,q} \right|, \text{ where } S_{k,g,i,q} \leq S_{k,g,i,q-1},$$

where $S_{k,g,i,q}$ is the number of shares held by investor k in firm i in quarter q classified into subgroup g and $P_{i,q}$ is the share price of firm i in quarter q. The investor's portfolio turnover for each subgroup is then defined as $PTR_{k,g,q} = \frac{\min\left(AgBuy_{k,g,q}, AgSell_{k,g,q}\right)}{\sum_{i \in N_{k,g}} (S_{k,g,i,q} P_{i,q} + S_{k,g,i,q-1} P_{i,q-1})/2}$.

The sample is limited to investors with available holding information in either subgroup, meaning that investors whose holdings are consistently ranked among the top five largest across all underlying securities, as well as those whose holdings are minor in all securities, are excluded. The final sample consists of 69,261 investor-quarter pairs and 138,522 observations, covering the

period from 1980 to 2022.

Panel A of Table 7 compares the distribution of portfolio turnover (PTR) between the Top5 subgroup and the Non-Top5 subgroup. Portfolio turnovers of Top5 are substantially smaller than those of Non-Top5 in every percentile. For instance, the median value of PTR is 0.230 in Non-Top5, nearly four times the median PTR in Top5. Even when the threshold is altered to the top 10 ranking, a significant discrepancy is still observed in PTR across the Top10 and Non-Top10 subgroups, as shown at the bottom of Panel A.

To mitigate the omitted variable concern, we estimate the following multi-variable regression model:

$$PTR_{k,g,q} = a + b_1 DumTop5_{k,g,q} + b_2 \chi_{k,g,q} + FE_{k,q} + \varepsilon_{k,g,q},$$

where $DumTop5$ equals 1 for the Top5 subgroup and 0 for the Non-Top5 subgroup. χ denotes a list of portfolio-level control variables: (i) PIO, the portfolio institution ownership calculated as the holding-weighted average of stock-level institution ownership, (ii) $PRet$, the portfolio quarterly return, (iii) $PRetStd$, the portfolio volatility calculated as the standard deviation of quarterly returns in the past two years, and (iv) $PSize$, the portfolio size computed as the logarithm of holding amount in million dollars. We also include investor-quarter fixed effects to account for trends in PTR that are investor-specific and may change over time. That said, the coefficient on $DumTop5$ should be interpreted as the within-investor-quarter difference in portfolio turnover between Top5 and Non-Top5.

Panel B of Table 7 reports the regression result. The coefficient on $DumTop5$ is significantly negative at -0.143 in Column (1), indicating that the portfolio turnover of Top5 is, on average, 14.3% lower than that of Non-Top5. As shown in Column (2), the results hold when we relax the threshold to the top 10 ranking.

Information content of earnings announcements. In the case of information shock, the price of security with more concentrated ownership is expected to reflect new information more

slowly because large investors refrain from trading aggressively. Hence, another hypothesis underlying the information pass-through channel is that the information content of the stock price surrounding an information shock is lower for firms with more concentrated active institutional ownership.

To test this hypothesis, we utilize quarterly earnings announcements that capture information shock. Following Landsman et al. (2012), we employ abnormal trading volume ($AVOL$) and abnormal return volatility ($AVAR$) to measure the information content of earnings announcements. $AVOL$ is calculated as the average trading volume within an event window, scaled by the counterparts within a non-event window: $AVOL = \ln\left(\frac{\overline{Volume}_{i,t\in[0,1]}}{\overline{Volume}_{i,t\in[-40,-6]}}\right)$, where $Volume_{i,t}$ denotes the daily trading volume in shares. Similarly, $AVAR$ is calculated as the mean square of adjusted returns within an event window, scaled by the counterparts within a non-event window: $AVAR = \ln\left(\frac{\overline{u^2}_{i,t\in[0,1]}}{\overline{u^2}_{i,t\in[-40,-6]}}\right)$, where $u_{i,t} = R_{i,t} - (\alpha_i + \beta_i R_{mkt,t})$ is calculated as daily stock returns subtracted by expected returns, with expected returns estimated against the market model across forty trading days before the announcement date to six trading days before the announcement date.

We apply the same rule introduced in Section C of the online appendix to pin down the effective earnings announcement date. Per Pevzner, Xie, and Xin (2015), we choose a two-day event window because newswire information is typically available on the next trading day. We commence the estimation window at $t-40$ to avoid overlapping the previous quarterly announcement date and conclude it at $t-6$ to prevent contaminating the parameter estimates with pre-leaked earnings information.

To investigate the effect of active institutional ownership concentration on the information content of earnings announcements, we then conduct the following regression model:

$$InfoContent_{i,q} = a + b_1 Concentration_{i,q} + b_2 \chi_{i,q} + FE_{i,q} + \varepsilon_{i,q},$$

where $InfoContent$ denotes the aforementioned two measures, $AVAR$ and $AVOL$; $\chi_{i,q}$ is the

same list of control variables as in the baseline regression model but on a quarterly basis; and $FE_{i,q}$ captures the firm fixed effect and quarter-industry two-way fixed effects.

Table 8 reports the results. In Panel A, the coefficient of interest, b_1, is significantly negative in all specifications, suggesting that less information is incorporated into the stock price for firms with more concentrated ownership. In Panel B, we further control for a saturated set of characteristics as in Pevzner, Xie, and Xin (2015) and demonstrate the notable negative relation between information content and active institutional ownership concentration.[12]

6.2 Learning Effect

The learning channel posits that the polarization of investor sizes impedes small investors from diversifying their learning, thus inducing them to allocate their learning capacity to a specific portfolio that favors assets with the largest supply. While large investors may diversify their learning, the impact can be limited since their learning is already well-diversified. Thus, a testable hypothesis is that greater concentration leads to increased learning in large stocks and diminished learning in small stocks.

To examine this hypothesis, we employ the download records of company filings from SEC EDGAR as an indicator of institutional investors' learning choices. We follow the data-cleaning process outlined by Ryans (2017) [13] and differentiate EDGAR downloads by active financial institutions from other market participants via two steps: First, we identify the IP addresses of active financial institutions by matching 13F active investors with two geolocation datasets, MaxMind and IPinfo, which provide information on IP addresses and their associated organizations. Sec-

[12]We include the following control variables, per Pevzner et al. (2015): $FirmSize$ denotes the natural logarithm of the market capitalization at the fiscal quarter end; $|UE|$ is the absolute value of unexpected earnings, computed as actual annual earnings minus the most recent median analyst forecast scaled by the quarter-end stock price; $ReportLag$ is the number of days from the fiscal quarter-end to the earnings announcement date; $ForeDisp$ is the standard deviation of analysts' earnings forecasts scaled by the fiscal quarter-end stock price; and $ForeNum$ is the number of annual earnings forecasts reported by LSEG IBES.

[13]First, we retain records of when the EDGAR server successfully delivered the requested document (code = 200). Second, we remove index page observations (idx = 1), as these pages provide the viewer a link to a filing, not the filing itself. Third, we exclude downloads identified as web crawlers by the EDGAR server (crawler = 1). Lastly, we filter out downloads by bots using three criteria: (1) an IP address downloads more than 25 items in one minute; (2) an IP address downloads items from more than three different companies in one minute; and (3) an IP address downloads more than 500 items in a single day.

ond, we identify active investors' use of EDGAR. We de-anonymize the partially obfuscated IP addresses in SEC EDGAR downloads using the mapping table from Chen et al. (2020), then match 13F active investors to EDGAR downloads based on these IP addresses.

To capture learning choices across various size groups, firms are categorized into five size groups based on their market capitalization and computed size-weighted average EDGAR downloads for each group in each quarter. To account for time-varying trends in overall learning capacity, downloads per group are normalized by the total downloads. The sample period spans from the first quarter of 2003 to the first quarter of 2017.

Panel (a) of Figure 13 illustrates the EDGAR downloads by active investors for each size group. Aligned with our theoretical implications, learning choices follow a water-filling pattern. Over 60% of download activities occur in the largest group compared to merely 4% in the smallest group. There is a concern that this pattern might primarily reflect the learning decisions of large investors rather than small ones. Due to strategic substitutability in learning, small investors may prefer to learn about assets that are different from those explored by large investors. If this effect is strong, small investors might allocate more learning capacity to smaller assets as ownership concentration increases. To attenuate this concern, Panels (b) and (c) separate downloads by large active investors (i.e., those with AUM above the sample median each quarter) and small active investors. The water-filling pattern exists in both investor groups, indicating that the strategic substitutability effect, if present, does not significantly alter small investors' learning choices.

Figure 14 investigates how size-based ownership concentration amongst active institutional investors impacts their learning choices. Panel (a1) identifies a significantly positive correlation between size-based $ActHHI_{size}$ and EDGAR downloads in the largest group. This suggests that investor attention is increasingly allocated to large stocks as size-based ownership concentration increases. The pattern reverses in the smallest group, as depicted in Panel (a2), meaning small stocks are poorly learned when active institutional ownership is concentrated. Meanwhile, Panel (a3) measures learning imbalance by calculating the difference in EDGAR downloads between the largest and smallest groups. Consistent with our hypotheses and numerical results in Figure

7, Panel (b), a positive correlation is observed between size-based ownership concentration and learning imbalance. The results remain robust when using $ActTop5_{size}$ as an alternative measure of ownership concentration, as demonstrated in Panels (b1) through (b3) of Figure 14.

7 Conclusion

Over the past several decades, equity ownership has increasingly become concentrated in the hands of large investors. As analyzed in the theoretical work of Kacperczyk et al. (2025), this skewed ownership structure can lower the informational efficiency of stock prices. In this paper, we provide the first empirical evidence for this negative effect of institutional ownership concentration on price informativeness.

To better inform our empirical analysis, we extend the theoretical framework of Kacperczyk et al. (2025) and conduct two novel analyses. First, we propose a new trading-based ownership concentration measure that leverages the rich and highly granular firm-level data on investor holdings, thereby enhancing the power of our empirical analysis. We find that an increase in Kacperczyk et al.'s (2025) size-based ownership concentration can lead to an increase in our trading-based ownership concentration. This implies a negative effect of trading-based ownership concentration on price informativeness, which we robustly support with extensive empirical evidence. Second, we explicitly differentiate between active and passive investors, revealing that it is the ownership concentration specifically among active investors that depresses price informativeness.

References

Agarwal, V., W. Jiang, Y. Tang, and B. Yang (2013). Uncovering hedge fund skill from the portfolio holdings they hide. *The Journal of Finance 68*(2), 739–783.

Agrawal, A. and Y. Lim (2022). Where do shareholder gains in hedge fund activism come from? Evidence from employee pension plans. *Journal of Financial and Quantitative Analysis 57*(6), 2140–2176.

Akey, P., V. Grégoire, and C. Martineau (2022). Price revelation from insider trading: Evidence from hacked earnings news. *Journal of Financial Economics 143*(3), 1162–1184.

Albuquerque, R., V. Fos, and E. Schroth (2022). Value creation in shareholder activism. *Journal of Financial Economics 145*(2), 153–178.

Antón, M., F. Ederer, M. Giné, and M. Schmalz (2023). Common ownership, competition, and top management incentives. *Journal of Political Economy 131*(5), 1294–1355.

Appel, I. R., T. A. Gormley, and D. B. Keim (2016). Passive investors, not passive owners. *Journal of Financial Economics 121*(1), 111–141.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2020). The fall of the labor share and the rise of superstar firms. *Quarterly Journal of Economics 135*(2), 645–709.

Bai, J., T. Philippon, and A. Savov (2016). Have financial markets become more informative? *Journal of Financial Economics 122*(3), 625–654.

Baker, A. C., D. F. Larcker, and C. C. Wang (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics 144*(2), 370–395.

Ball, R. and P. Brown (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research 6*(2), 159–178.

Ben-David, I., F. Franzoni, R. Moussawi, and J. Sedunov (2021). The granular nature of large institutional investors. *Management Science 67*(11), 6629–6659.

Bennett, B., R. M. Stulz, and Z. Wang (2020). Does joining the S&P 500 index hurt firms? Technical report, National Bureau of Economic Research.

Boehmer, E. and E. K. Kelley (2009). Institutional investors and the informational efficiency of prices. *Review of Financial Studies 22*(9), 3563–3594.

Bogousslavsky, V., V. Fos, and D. Muravyev (2024). Informed trading intensity. *Journal of Finance 79*(2), 903–948.

Bond, P., A. Edmans, and I. Goldstein (2012). The real effects of financial markets. *Annual Review of Financial Economics 4*(1), 339–360.

Boone, A. L. and J. T. White (2015). The effect of institutional ownership on firm transparency and information production. *Journal of Financial Economics 117*(3), 508–533.

Bushee, B. J. (1998). The influence of institutional investors in myopic R&D investment behavior. *Accounting Review 73*(3), 305–333.

Cahan, S. F., C. Chen, and L. Chen (2024). In financial statements we trust: institutional investors' stockholdings after restatements. *The Accounting Review 99*(2), 143–168.

Cai, C. X., K. Keasey, and H. Short (2006). Corporate governance and information efficiency in security markets. *European Financial Management 12*(5), 763–787.

Cao, C., B. Liang, A. W. Lo, and L. Petrasek (2018). Hedge fund holdings and stock market efficiency. *The Review of Asset Pricing Studies 8*(1), 77–116.

Cao, J., S. Titman, X. Zhan, and W. Zhang (2023). ESG preference, institutional trading, and stock return patterns. *Journal of Financial and Quantitative Analysis 58*(5), 1843–1877.

Cen, L., B. Han, Y. Han, and C. Jo (2024). Data scientists on Wall Street. *Working Paper*.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The effect of minimum wages on low-wage jobs. *Quarterly Journal of Economics 134*(3), 1405–1454.

Chen, H., L. Cohen, U. Gurun, D. Lou, and C. Malloy (2020). IQ from IP: Simplifying search in portfolio choice. *Journal of Financial Economics 138*(1), 118–137.

Chen, S., X. Chen, and Q. Cheng (2008). Do family firms provide more or less voluntary disclosure? *Journal of Accounting Research 46*(3), 499–536.

Chinco, A. and M. Sammon (2024). The passive ownership share is double what you think it is. *Journal of Financial Economics 157*, 103860.

Coles, J. L., D. Heath, and M. C. Ringgenberg (2022). On index investing. *Journal of Financial Economics 145*(3), 665–683.

Dávila, E. and C. Parlatore (2025). Identifying price informativeness. *The Review of Financial Studies*, hhaf051.

Dechow, P. M., W. Ge, C. R. Larson, and R. G. Sloan (2011). Predicting material accounting misstatements. *Contemporary Accounting Research 28*(1), 17–82.

Dong, X., N. Kang, and J. Peress (2025). Fast and slow arbitrage: The predictive power of (persistent) capital flows for factor returns. *The Review of Financial Studies*, hhaf036.

Duarte, J., E. Hu, and L. Young (2020). A comparison of some structural models of private information arrival. *Journal of Financial Economics 135*(3), 795–815.

Duarte, J. and L. Young (2009). Why is PIN priced? *Journal of Financial Economics 91*(2), 119–138.

Dyck, A., A. Morse, and L. Zingales (2010). Who blows the whistle on corporate fraud? *The Journal of Finance 65*(6), 2213–2253.

Easley, D., N. M. Kiefer, M. O'hara, and J. B. Paperman (1996). Liquidity, information, and infrequently traded stocks. *Journal of Finance 51*(4), 1405–1436.

Edmans, A., S. Jayaraman, and J. Schneemeier (2017). The source of information in prices and investment-price sensitivity. *Journal of Financial Economics 126*(1), 74–96.

Eisfeldt, A. L. and D. Papanikolaou (AER2014). The value and ownership of intangible capital. *American Economic Review 104*(5), 189–194.

Farboodi, M., A. Matray, L. Veldkamp, and V. Venkateswaran (2022). Where has all the data gone? *Review of Financial Studies 35*(7), 3101–3138.

Friberg, R., I. Goldstein, and K. W. Hankins (2024). Corporate responses to stock price fragility. *Journal of Financial Economics 153*, 103795.

Glebkin, S., S. Malamud, and A. Mokak Teguia (2025). Benign granularity in asset markets.

Goldstein, I. (2023). Information in financial markets and its real effects. *Review of Finance 27*(1), 1–32.

Greenwood, R. and D. Thesmar (2011). Stock price fragility. *Journal of Financial Economics 102*(3),

471–490.

Guo, F., C. Lin, A. Masli, and M. S. Wilkins (2021). Auditor responses to shareholder activism. *Contemporary Accounting Research 38*(1), 63–95.

Hartzell, J. C. and L. T. Starks (2003). Institutional investors and executive compensation. *The Journal of Finance 58*(6), 2351–2374.

He, J. and J. Huang (2017). Product market competition in a world of cross-ownership: Evidence from institutional blockholdings. *Review of Financial Studies 30*(8), 2674–2718.

Hollingsworth, A., K. Karbownik, M. A. Thomasson, and A. Wray (2024). The gift of a lifetime: The hospital, modern medicine, and mortality. *American Economic Review 114*(7), 2201–2238.

Huang, J.-Z., Y. Wang, and Y. Wang (2024). Does ownership concentration affect corporate bond volatility? Evidence from bond mutual funds. *Journal of Banking & Finance 165*, 107217.

Ivković, Z., C. Sialm, and S. Weisbenner (2008). Portfolio concentration and the performance of individual investors. *Journal of Financial and Quantitative Analysis 43*(3), 613–655.

Jha, M., J. Qian, M. Weber, and B. Yang (2024). ChatGPT and corporate policies. Technical report, National Bureau of Economic Research.

Kacperczyk, M., J. Nosal, and S. Sundaresan (2025). Market power and price informativeness. *Review of Economic Studies 92*(3), 1955–1986.

Kacperczyk, M., C. Sialm, and L. Zheng (2005). On the industry concentration of actively managed equity mutual funds. *The Journal of Finance 60*(4), 1983–2011.

Kacperczyk, M., S. Sundaresan, and T. Wang (2021). Do foreign institutional investors improve price efficiency? *Review of Financial Studies 34*(3), 1317–1367.

Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp (2016). A rational theory of mutual funds' attention allocation. *Econometrica 84*(2), 571–626.

Kempf, E. and O. Spalt (2023). Attracting the sharks: Corporate innovation and securities class action lawsuits. *Management Science 69*(3), 1805–1834.

Kini, O., S. Lee, and M. Shen (2024). Common institutional ownership and product market threats. *Management Science 70*(5), 2705–2731.

Kwon, S. Y., Y. Ma, and K. Zimmermann (2024). 100 years of rising corporate concentration. *American Economic Review 114*(7), 2111–2140.

Kyle, A. S. (1989). Informed speculation with imperfect competition. *Review of Economic Studies 56*(3), 317–355.

Landsman, W. R., E. L. Maydew, and J. R. Thornock (2012). The information content of annual earnings announcements and mandatory adoption of IFRS. *Journal of Accounting and Economics 53*(1-2), 34–54.

Lee, C., K. H. Chung, and S. Yang (2016). Corporate governance and the informational efficiency of prices. *Financial Management 45*(1), 239–260.

Levonyan, V. and P. Mengano (2024). Mergers and market power. *Available at SSRN 4466826*.

Lewellen, K. and M. Lowry (2021). Does common ownership really increase firm coordination? *Journal of Financial Economics 141*(1), 322–344.

Liu, X. (2016). Corruption culture and corporate misconduct. *Journal of Financial Eco-*

nomics 122(2), 307–327.

Lo, A. W. and A. C. MacKinlay (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies 1*(1), 41–66.

Martineau, C. (2022). Rest in peace post-earnings announcement drift. *Critical Finance Review 11*(3-4), 613–646.

Massa, M., D. Schumacher, and Y. Wang (2021). Who is afraid of BlackRock? *Review of Financial Studies 34*(4), 1987–2044.

Muravyev, D., N. D. Pearson, and J. M. Pollet (2022). Is there a risk premium in the stock lending market? Evidence from equity options. *The Journal of Finance 77*(3), 1787–1828.

Odders-White, E. R. and M. J. Ready (2008). The probability and magnitude of information events. *Journal of Financial Economics 87*(1), 227–248.

Peress, J. (2010). Product market competition, insider trading, and stock market efficiency. *The Journal of Finance 65*(1), 1–43.

Peters, R. H. and L. A. Taylor (2017). Intangible capital and the investment-Q relation. *Journal of Financial Economics 123*(2), 251–272.

Pevzner, M., F. Xie, and X. Xin (2015). When firms talk, do investors listen? The role of trust in stock market reactions to corporate earnings announcements. *Journal of Financial Economics 117*(1), 190–223.

Porras Prado, M., P. A. Saffi, and J. Sturgess (2016). Ownership structure, limits to arbitrage, and stock returns: Evidence from equity lending markets. *Review of Financial Studies 29*(12), 3211–3244.

Rogers, J. L. and P. C. Stocken (2005). Credibility of management forecasts. *The Accounting Review 80*(4), 1233–1260.

Ryans, J. (2017). Using the edgar log file data set. *Available at SSRN 2913612*.

Sammon, M. (2024). Passive ownership and price informativeness. *Management Science*.

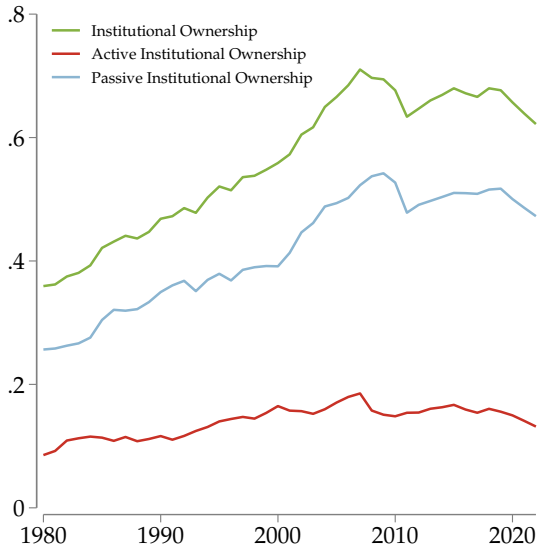
Sialm, C. and D. X. Xu (2023). Information acquisition by mutual fund investors: Evidence from stock trading suspensions. *Available at SSRN 5047094*.

Subrahmanyam, A. and S. Titman (2001). Feedback from stock prices to cash flows. *Journal of Finance 56*(6), 2389–2413.

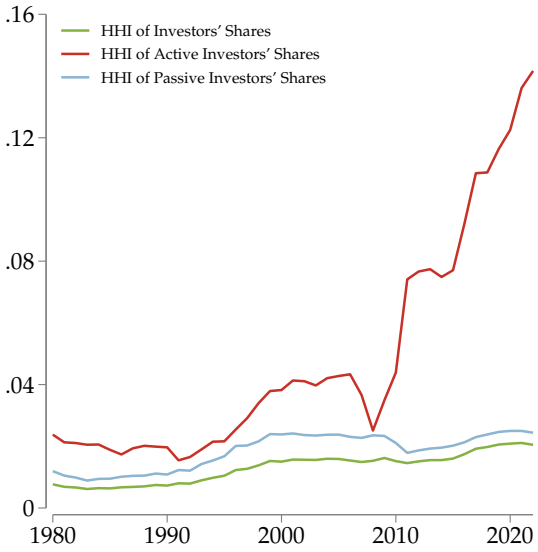
Wang, W. (2022). Accounting restatements and bank liquidity creation. *The Accounting Review 97*(6), 445–473.

Wardlaw, M. (2020). Measuring mutual fund flow pressure as shock to stock returns. *The Journal of Finance 75*(6), 3221–3243.

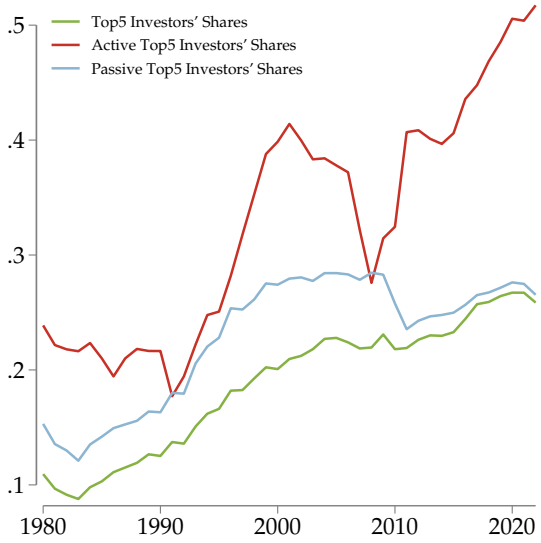
Yan, X. and Z. Zhang (2009). Institutional investors and equity returns: Are short-term institutions better informed? *Review of Financial Studies 22*(2), 893–924.



(a) Ownership



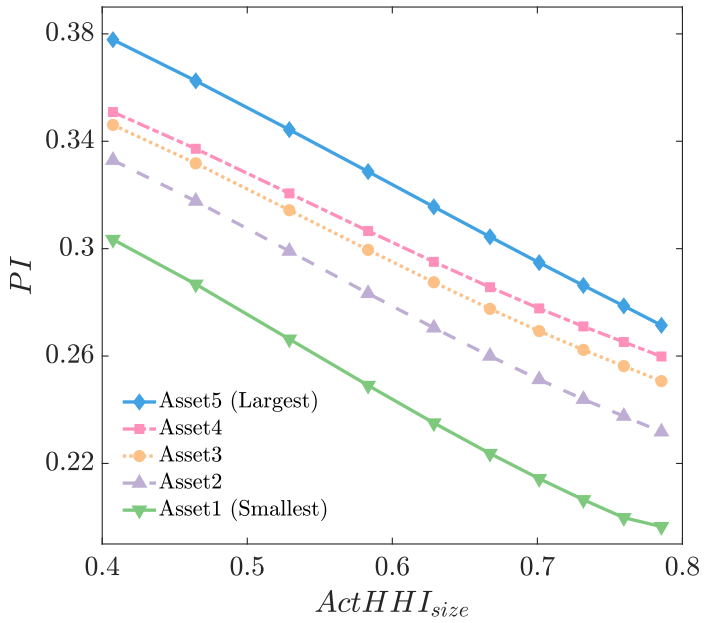
(b) Concentration: HHI



(c) Concentration: Top 5 Investor Shares

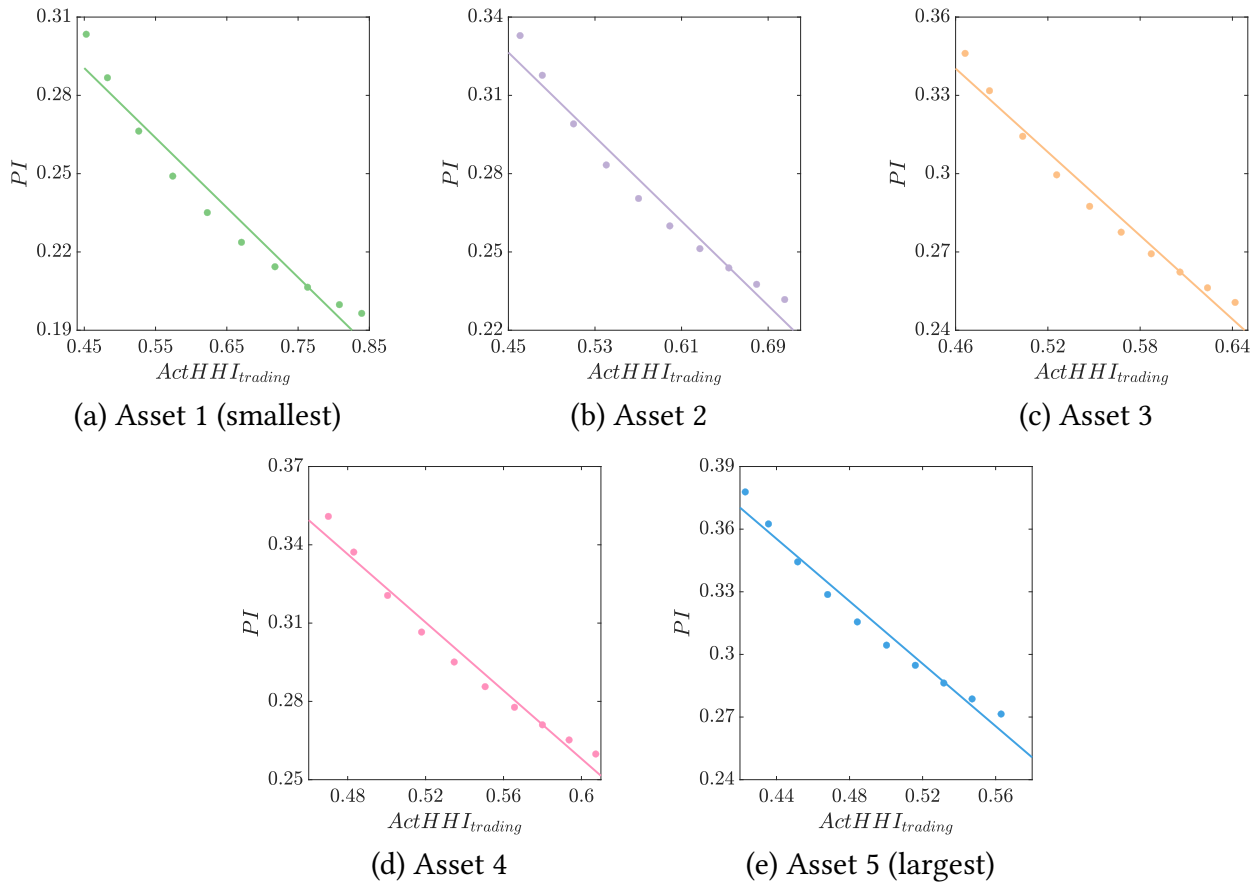
Panel (a) of this figure plots the total institutional equity ownership as well as the breakdown into active and passive ownership. Panels (b) and (c) present measures of the concentration of institutional investors within each group, specifically the Herfindahl-Hirschman Index (HHI) of investors' assets under management (AUM) and the share of AUM held by the top five investors. Investor AUM is calculated as the aggregate value of their equity holdings.

Figure 1: The Time Trend of Institutional Ownership and Its Concentration



This figure plots the individual price informativeness against different values of size-based ownership concentration. Price informativeness and size-based ownership concentration are defined in equations (6) and (7) respectively. The individual assets are ranked by their supply $\bar{x}$, from the smallest (asset 1) to the largest (asset 5).

Figure 2: The Effect of Size-based Ownership Concentration

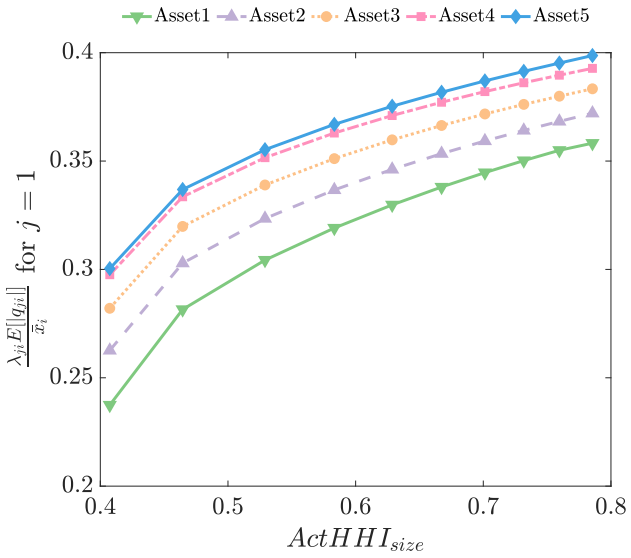


(a) Asset 1 (smallest) (b) Asset 2 (c) Asset 3

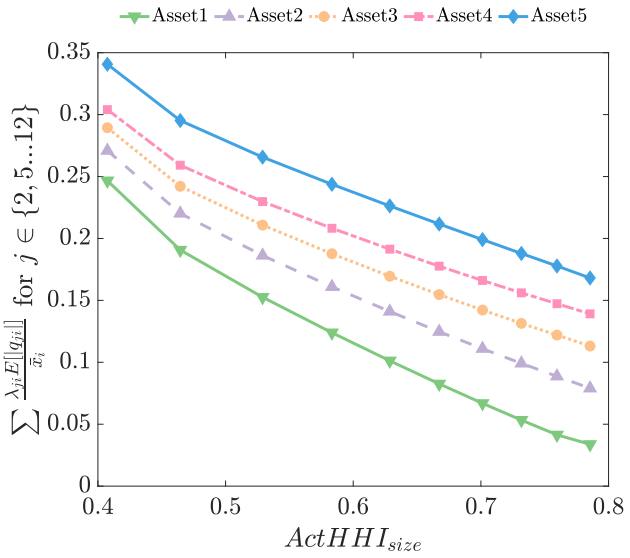
(d) Asset 4 (e) Asset 5 (largest)

This figure plots individual price informativeness against different values of trading-based ownership concentration at the asset level. Price informativeness and trading-based ownership concentration are defined in equations (6) and (8) respectively. The individual assets are ranked by their supply $\bar{x}$, from the smallest (asset 1) to the largest (asset 5).

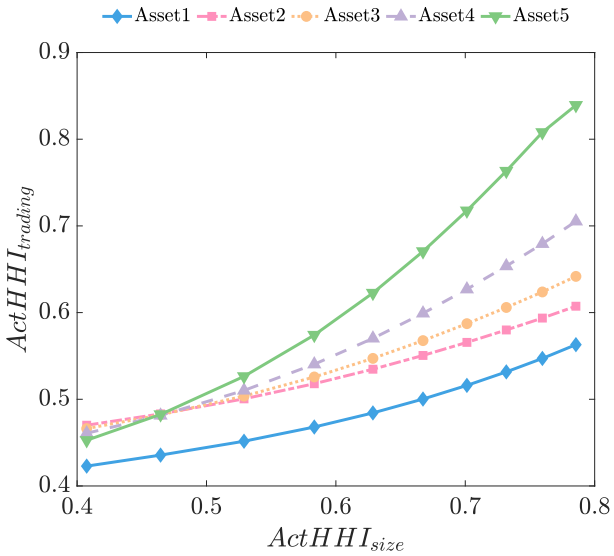
Figure 3: The Effect of Trading-based Ownership Concentration



(a) Trading Percentage by the Largest Active Investors



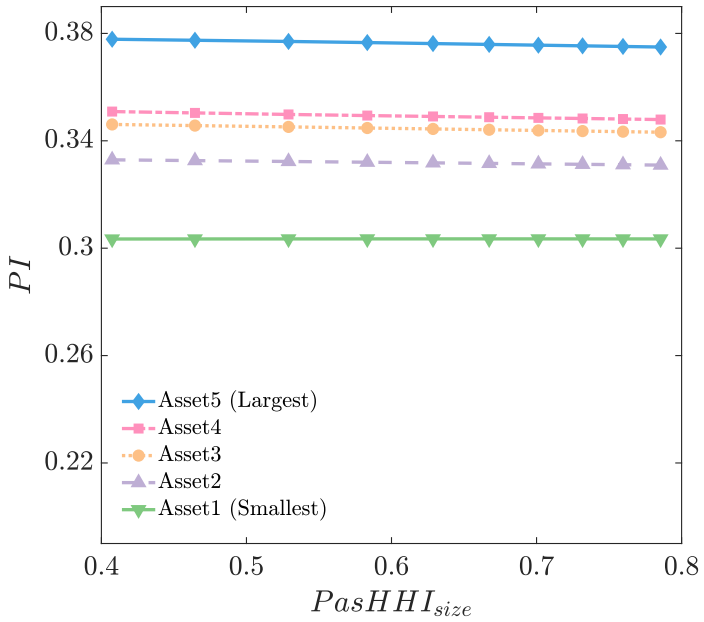
(b) Trading Percentage by Other Active Investors



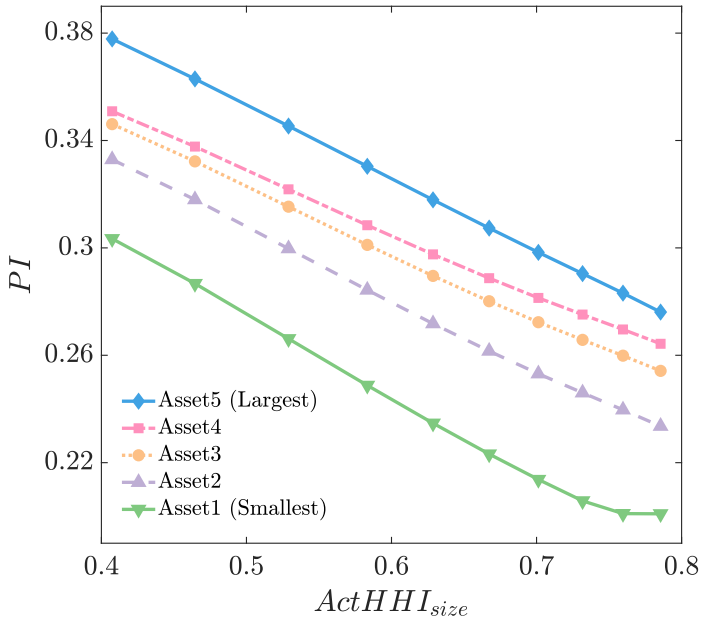
(c) $ActHHI_{size}$ and $ActHHI_{trading}$

Panels (a) and (b) plot active investors' trading choices against different values of size-based ownership concentration. Panel (c) plots the trading-based ownership concentration values against the size-based ownership concentration values. The individual assets are ranked by their supply $\bar{x}$, from the smallest (asset 1) to the largest (asset 5).

Figure 4: Trading Choices by Active Investors



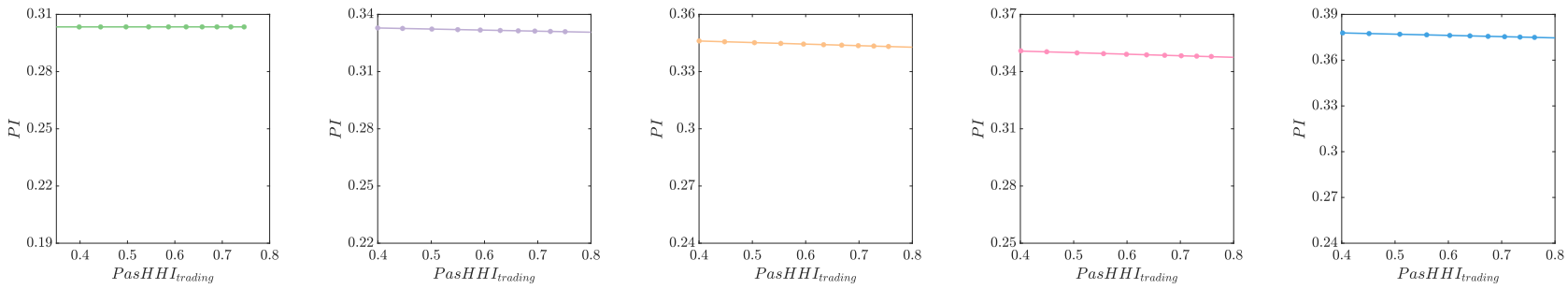
(a) Sized-based Ownership Concentration among Passive Investors



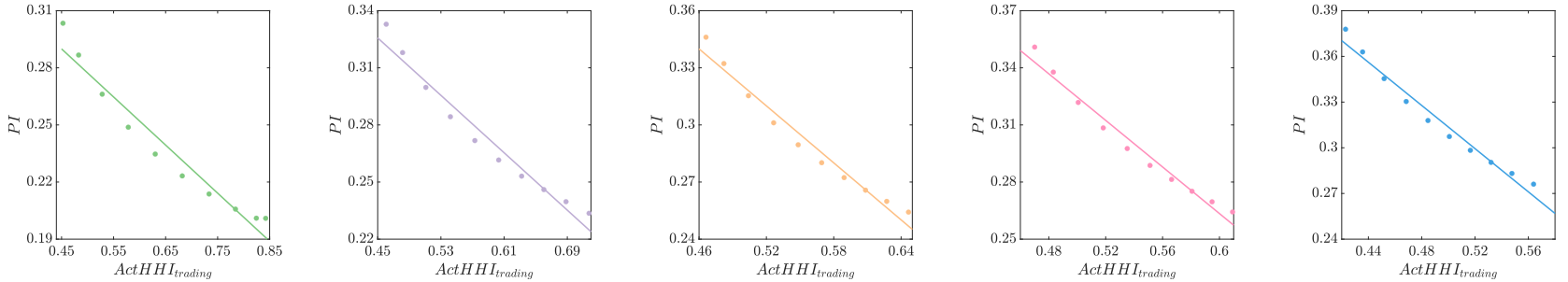
(b) Sized-based Ownership Concentration among Active Investors

This figure illustrates the effect of size-based ownership concentration among passive investors in Panel (a) and that among active investors in Panel (b). In Panel (a), the size distribution of active investors is fixed ($\lambda_1/\lambda_2 = 1.1$ and $\sum_{j\in SA}\lambda_j/\sum_{j\in SA\cup LA}\lambda_j = 0.10$), while in Panel (b), the size distribution of passive investors is fixed ($\lambda_3/\lambda_4 = 1.1$ and $\sum_{j\in SP}\lambda_j/\sum_{j\in SP\cup LP}\lambda_j = 0.10$).

Figure 5: Size-based Ownership Concentration among Passive/Active Investors



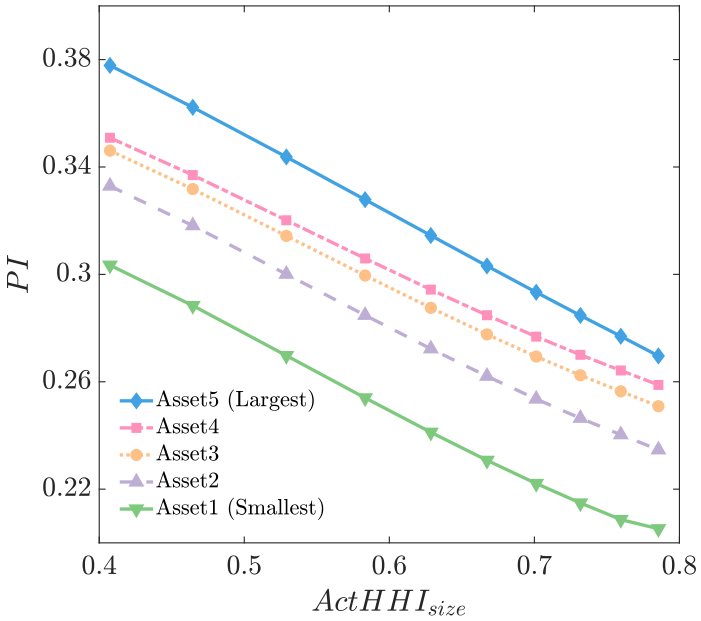
(a1) Asset 1 (smallest) (a2) Asset 2 (a3) Asset 3 (a4) Asset 4 (a5) Asset 5 (largest)



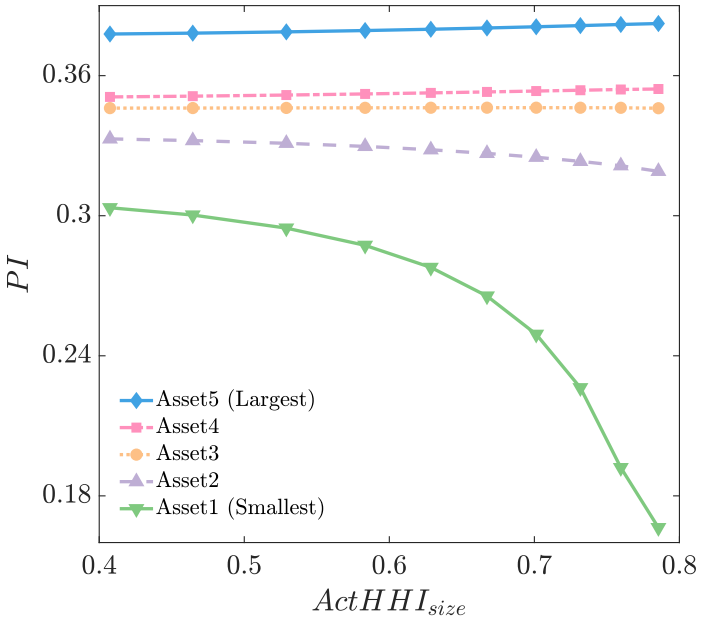
(b1) Asset 1 (smallest) (b2) Asset 2 (b3) Asset 3 (b4) Asset 4 (b5) Asset 5 (largest)

This figure illustrates the effect of trading-based ownership concentration among passive investors in Panels (a1)-(a5) and that among active investors in Panels (b1)-(b5). In Panels (a1)-(a5), the size distribution of active investors is fixed ($\lambda_1/\lambda_2 = 1.1$ and $\sum_{j\in SA}\lambda_j/\sum_{j\in SA\cup LA}\lambda_j = 0.10$), while in Panels (b1)-(b5), the size distribution of passive investors is fixed ($\lambda_3/\lambda_4 = 1.1$ and $\sum_{j\in SP}\lambda_j/\sum_{j\in SP\cup LP}\lambda_j = 0.10$).

Figure 6: Trading-based Ownership Concentration among Passive/Active Investors



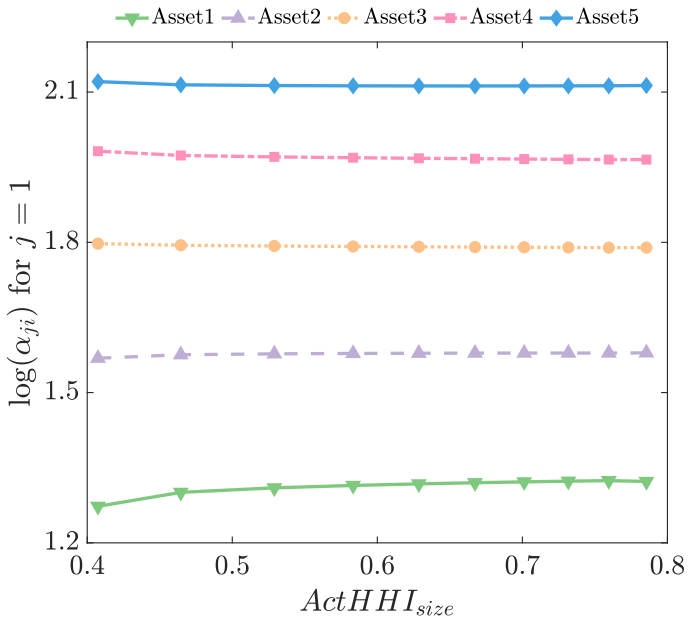
(a) Information Pass-through Effect



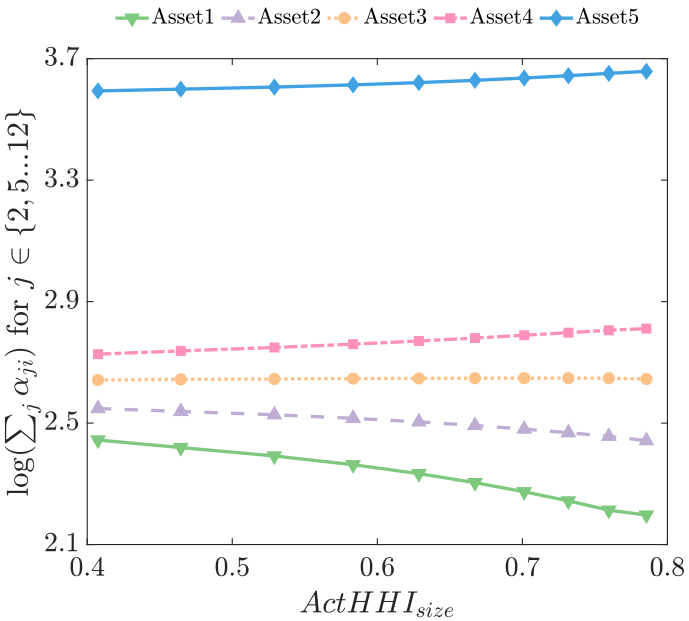
(b) Learning Effect

Panels (a) and (b) decompose the overall effect of ownership concentration documented in Figure 2 by respectively fixing the degree of learning (α_{ji}) and fixing the information pass-through (ω_{ji}). Price informativeness and size-based ownership concentration are defined in equations (6) and (7) respectively. The individual assets are ranked by their supply $\bar{x}$, from the smallest (asset 1) to the largest (asset 5).

Figure 7: Decomposition of the Effect of Ownership Concentration



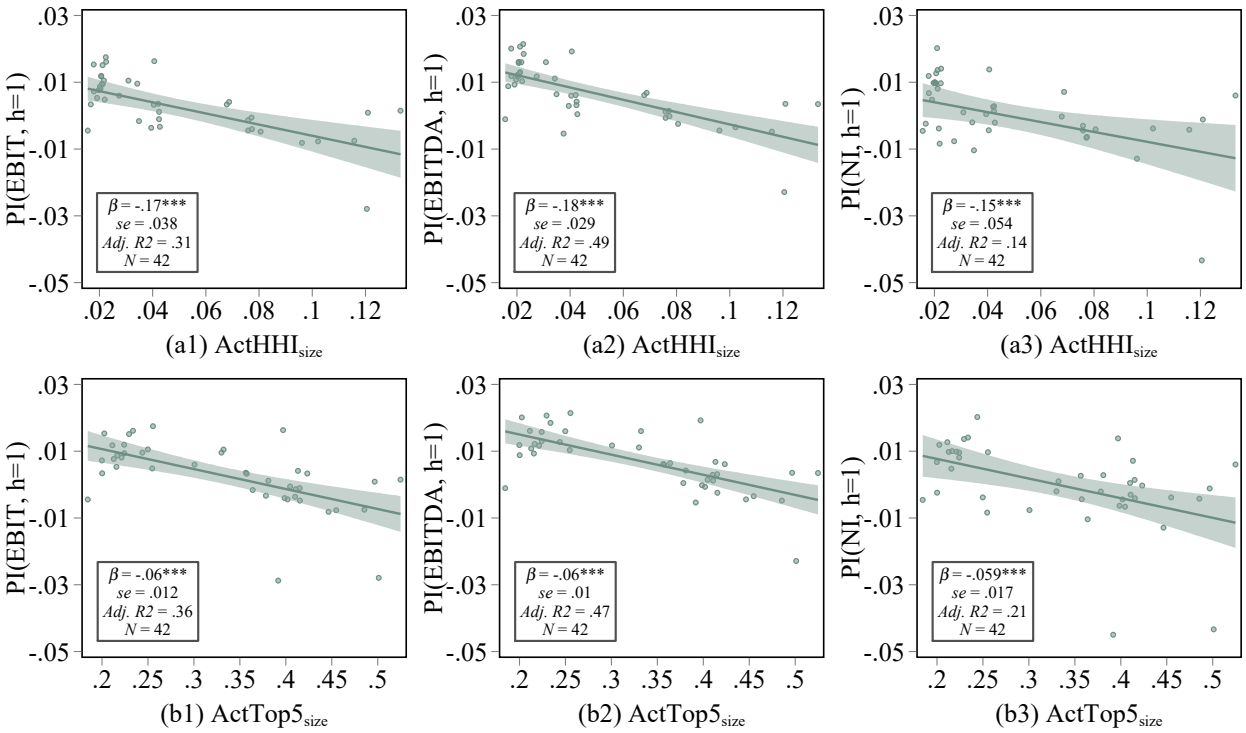
(a) Learning by the Largest Active Investors



(b) Learning by All Other Active Investors

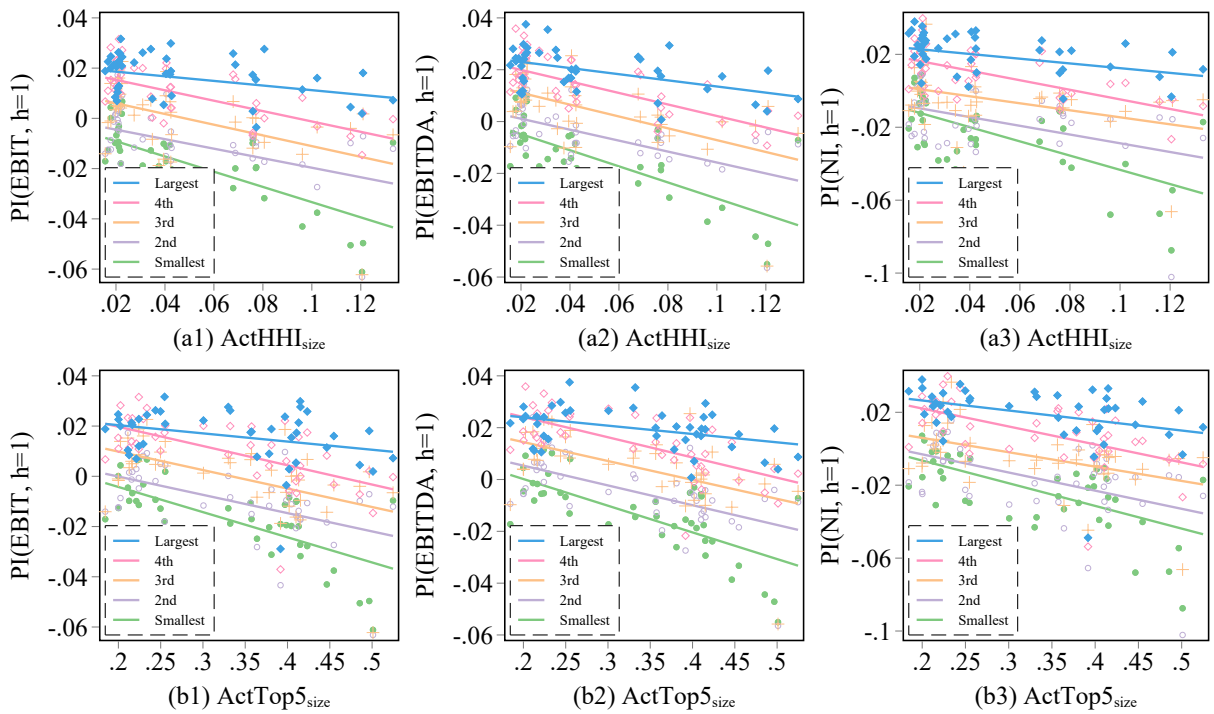
This figure plots active investors' learning choices against different values of size-based ownership concentration. The individual assets are ranked by their supply $\bar{x}$, from the smallest (asset 1) to the largest (asset 5).

Figure 8: Learning Choices by Active Investors



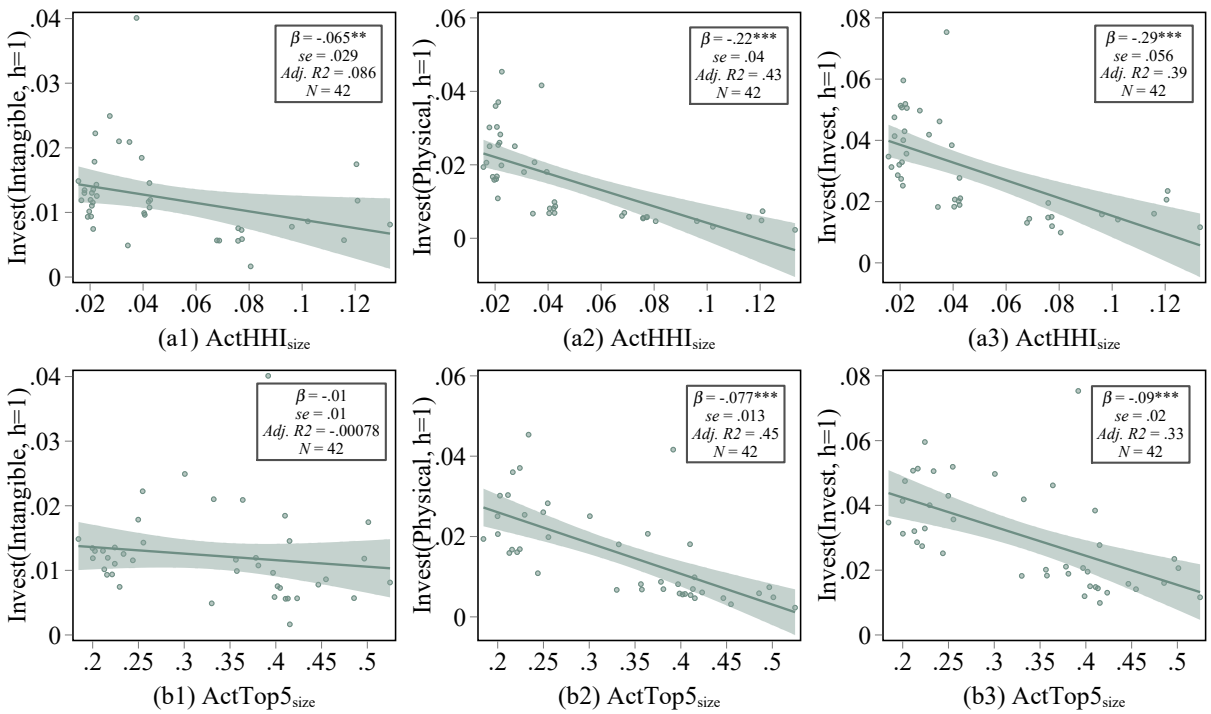
This figure presents scatter plots illustrating the relation between price informativeness and size-based ownership concentration. The plots include fitted lines and 95% confidence intervals. Size-based ownership concentration is quantified using two metrics: (i) $ActHHI_{size}$: the Herfindahl-Hirschman Index of AUM among active institutional investors, shown in Panels (a1)-(a3), and (ii) $ActTop5_{size}$: the proportion of AUM held by the top five active institutional investors relative to the total AUM of all active institutional investors, depicted in Panels (b1)-(b3). PI is derived from equations (12) and (13) and measures the predictability of future cash flows based on current market prices, with future cash flows represented by one of the three variables ($EBIT$, $EBITDA$, or NI) calculated as of year $t + h$, scaled by total assets in year t. The prediction horizon, denoted by h, is set at 1 year. The sample has an annual frequency and spans from 1980 to 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Figure 9: Price Informativeness and Size-based Ownership Concentration



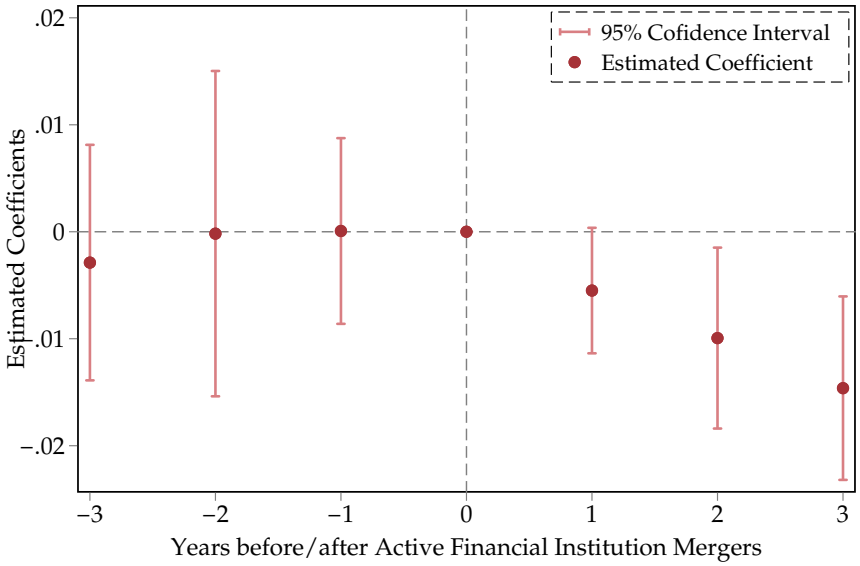
This figure presents scatter plots with fitted lines illustrating the relation between price informativeness by size group and size-based ownership concentration. Size-based ownership concentration is quantified using two metrics: (i) $ActHHI_{size}$: the Herfindahl-Hirschman Index of AUM among active institutional investors, shown in Panels (a1)-(a3), and (ii) $ActTop5_{size}$: the proportion of AUM held by the top five active institutional investors relative to the total AUM of all active institutional investors, depicted in Panels (b1)-(b3). We divide the sample firms into quintiles based on each security's market capitalization, and estimate PI for each group according to equations (12) and (13). Future cash flows in equation (12) are represented by one of the three variables ($EBIT$, $EBITDA$, or NI) calculated as of year $t+h$, scaled by total assets in year t. The prediction horizon, denoted by h, is set at 1 year. The sample has an annual frequency and spans from 1980 to 2022.

Figure 10: Price Informativeness by Size Group and Size-based Ownership Concentration

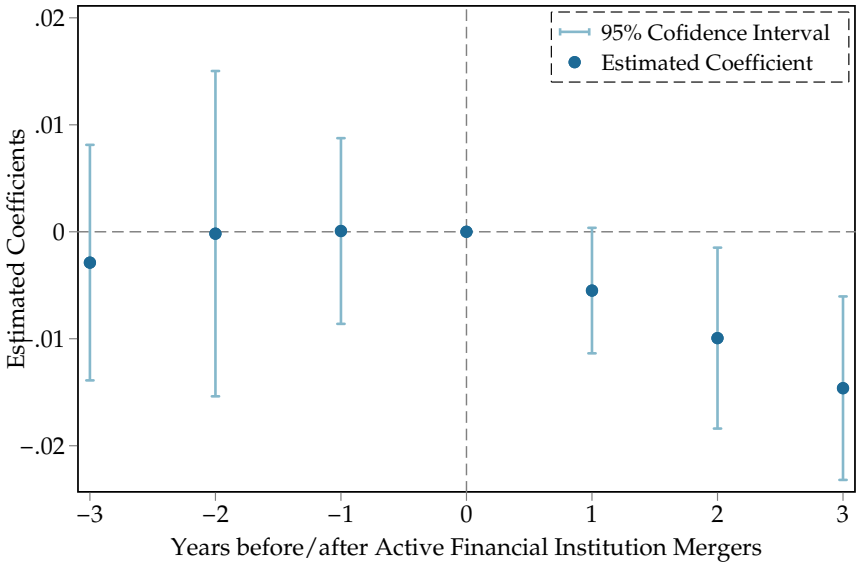


This figure presents scatter plots illustrating the relation between investment-to-price sensitivity and size-based ownership concentration. The plots include fitted lines and 95% confidence intervals. Size-based ownership concentration is quantified using two metrics: (i) $ActHHI_{size}$: the Herfindahl-Hirschman Index of AUM among active institutional investors, shown in Panels (a1)-(a3), and (ii) $ActTop5_{size}$: the proportion of AUM held by the top five active institutional investors relative to the total AUM of all active institutional investors, depicted in Panels (b1)-(b3). Investment-to-price sensitivity is derived from equation (14) and measures the extent to which current market prices reveal the information necessary for future investment decisions, with future investments represented by one of the three variables ($Intangible$, $Physical$, or $Invest$) calculated as of year $t + h$, scaled by total capital in year t. The prediction horizon, denoted by h, is set at 1 year. The sample has an annual frequency and spans from 1980 to 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Figure 11: Investment-to-Price Sensitivity and Size-based Ownership Concentration



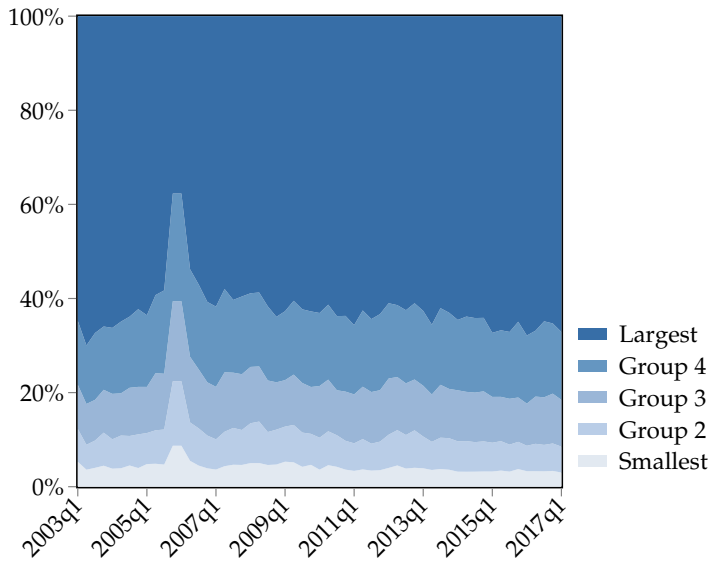
(a) Price Informativeness



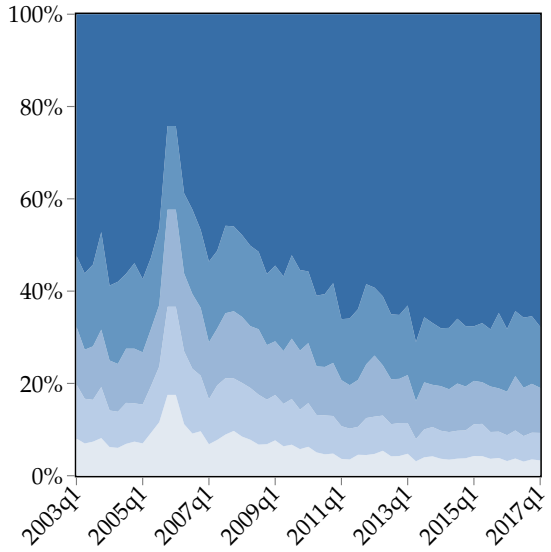
(b) Investment-to-Price Sensitivity

This figure plots the estimated coefficients on triple interactions of the market price variable ($\log(M/A)$) with treatment indicator variable ($Treat$) with a set of year dummy variables. The estimation window spans $(-3, +3)$ years, with year 0 denoting the merger completion year. Panel (a) measures price informativeness based on the earnings variable $EBITDA/A$ at the 1-year prediction horizon, while Panel (b) measures investment-to-price sensitivity based on the investment variable $Invest/K$ at the 1-year prediction horizon. We drop the interaction for the merger completion year (year 0) to avoid multicollinearity, and thus the effect is normalized to zero for that year. Standards errors are clustered at the year and firm levels.

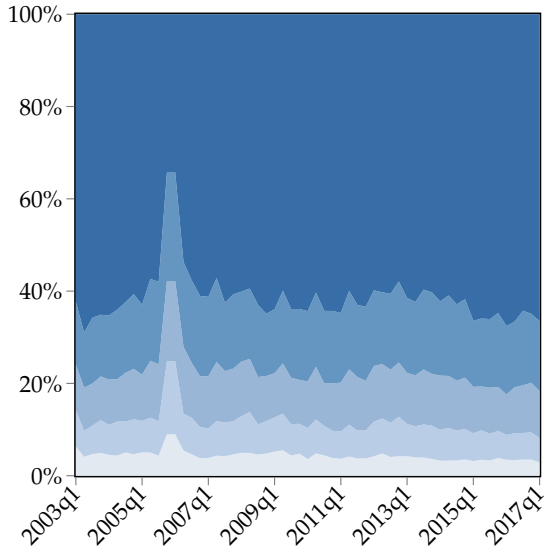
Figure 12: Event-Study Estimates



(a) All Active Investors



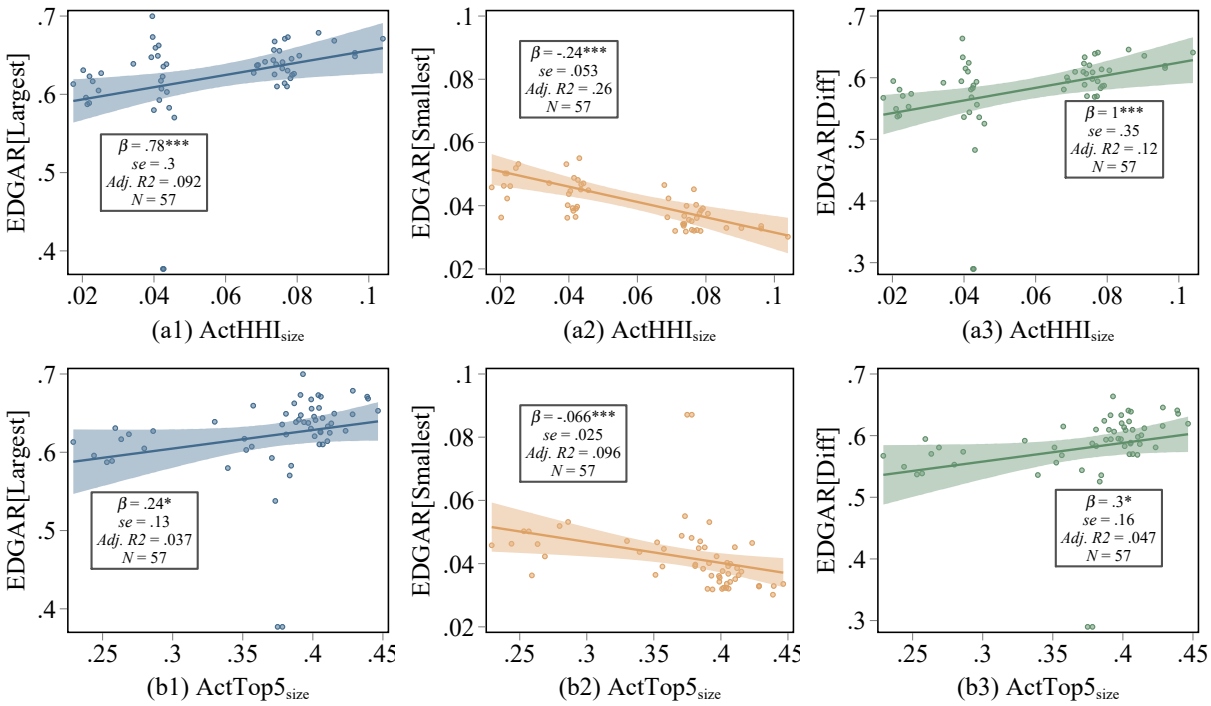
(b) Large Active Investors



(c) Small Active Investors

This figure depicts the time-series EDGAR downloads for each stock size group. The downloads within each stock size group are normalized by the total number of downloads, such that their sum equals one. The largest group comprises sample firms with the highest market capitalization at each quarter's end, while the smallest group includes those with the lowest market capitalization. Panel (a) considers EDGAR downloads from all active investors, while Panels (b) and (c) further separate downloads by large and small active investors.

Figure 13: EDGAR Downloads for Each Group



This figure presents scatter plots illustrating the relation between EDGAR downloads (in percentage) by size group and size-based ownership concentration, as measured by $ActHHI_{size}$ in Panels (a1)-(a3) and $ActTop5_{size}$ in Panels (b1)-(b3). Each plot includes fitted lines and 95% confidence intervals. Panels (a1) and (b1) focus on the weighted average EDGAR downloads in the group with the largest market capitalization, while Panels (a2) and (b2) focus on the weighted average EDGAR downloads in the group with the smallest market capitalization. Panels (a3) and (b3) examine the learning imbalance, defined as the difference in the weighted average EDGAR downloads between the largest and smallest groups. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Figure 14: Size-based Ownership Concentration and EDGAR Downloads

Table 1: Parameter Values in Numerical Analysis

Parameter	Symbol	Value
Mean payoff	$\bar{z}_i$	10
Supply	$\bar{x}_i$	$\in [3, 6]$, linear distribution across i
Number of assets, oligopolists	n, l	5, 20
Gross risk-free rate	r	1.025
Vol. of asset supply	$\sigma_{x,i}$	Coefficient of variation of 0.2 for all i
Vol. of asset payoffs	σ_i	1 for all i
Risk aversion	ρ	2.32
Information capacities	K_j	12.5 for $j \in LA$ and 1.25 for $j \in SA$
Retail investors	λ_0	0.4
Passive investors	$\frac{\sum_{j\in LP\cup SP}\lambda_j}{\sum_{j=1}^{l}\lambda_j}$	0.5
Small active investors	$\frac{\sum_{j\in SA}\lambda_j}{\sum_{j\in SA\cup LA}\lambda_j}$	Varying linearly from 0.10 to 0.03
Small passive investors	$\frac{\sum_{j\in SP}\lambda_j}{\sum_{j\in SP\cup LP}\lambda_j}$	Varying linearly from 0.10 to 0.03
Relative size within large investors	$\frac{\lambda_1}{\lambda_2}, \frac{\lambda_3}{\lambda_4}$	Varying linearly from 1.1 to 10
Relative size within small investors	$\frac{\lambda_5}{\lambda_{12}}, \frac{\lambda_{13}}{\lambda_{20}}$	5

Table 2: Summary Statistics

This table presents the summary statistics for the variables used in our main analysis. The sample has an annual frequency and spans from 1980 to 2022. All continuous variables are winsorized at the top and bottom 1% to mitigate the influence of outliers. Variable definitions are provided in Table B1 of the online appendix.

Variable	N	Mean	SD	p10	p25	p50	p75	p90
Panel A: Ownership Concentration Variables								
$ActHHI_{size}$	42	0.048	0.034	0.019	0.021	0.036	0.076	0.102
$ActTop5_{size}$	42	0.333	0.102	0.211	0.224	0.357	0.411	0.455
$ActHHI$	89218	0.239	0.163	0.080	0.120	0.196	0.309	0.457
$ActTop5$	89218	0.768	0.171	0.524	0.637	0.781	0.924	0.992
Panel B: Earning Variables								
$EBIT/A$	88269	0.048	0.178	-0.110	0.027	0.079	0.130	0.189
$EBITDA/A$	89114	0.092	0.177	-0.060	0.067	0.121	0.175	0.237
NI/A	89218	0.001	0.189	-0.157	-0.004	0.042	0.081	0.126
Panel C: Investment Rate Variables								
$Intangible/K$	88833	0.106	0.095	0.004	0.032	0.087	0.150	0.234
$Physical/K$	88286	0.063	0.066	0.011	0.022	0.043	0.078	0.137
$Invest/K$	88797	0.170	0.111	0.058	0.091	0.143	0.215	0.319
Panel D: Control Variables								
$log(M/A)$	89218	0.020	0.979	-1.192	-0.616	0.020	0.660	1.268
$PasHHI$	89218	0.586	0.190	0.360	0.439	0.556	0.721	0.875
$PasTop5$	89218	0.123	0.109	0.041	0.057	0.087	0.147	0.245
IO	89218	0.567	0.271	0.197	0.348	0.568	0.790	0.928
$Leverage$	89218	0.217	0.184	0.000	0.037	0.199	0.346	0.471
$Sale$	89218	1.058	0.749	0.281	0.526	0.917	1.387	1.979
$Cash$	89218	0.188	0.222	0.008	0.026	0.095	0.268	0.536

Table 3: Price Informativeness and Trading-based Ownership Concentration

This table reports OLS estimates on the relation between price informativeness, which measures the predictability of future cash flows from current market prices, and trading-based ownership concentration. The dependent variable is future earnings, calculated as one of the three cash flow variables ($EBIT$, $EBITDA$, and NI) in year $t+h$ divivded by total assets in year t. Here, h denotes the prediction horizons, set at 1 in Columns (1)-(3) and 3 in Columns (4)-(6). Concentration is measured by $ActHHI$ in Panel A, representing the Herfindahl-Hirschman index of active institutional ownership, and by $ActTop5$ in Panel B, denoting the proportion of shares held by the top five active institutional investors relative to the total shares held by all active institutional investors. $\log(M/A)$ is the log-ratio of a firm's market capitalization to its total assets. See Table B1 of the online appendix for the complete list of variable definitions. Standard errors, clustered at the year and firm levels, are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: $ActHHI$						
where E =	$E_{h=1}/A$ $EBIT$	$E_{h=1}/A$ $EBITDA$	$E_{h=1}/A$ NI	$E_{h=3}/A$ $EBIT$	$E_{h=3}/A$ $EBITDA$	$E_{h=3}/A$ NI
$\log(M/A)$	0.030***	0.041***	0.041***	0.013	0.048***	-0.028
	(0.011)	(0.007)	(0.013)	(0.011)	(0.010)	(0.018)
E/A	0.539***	0.559***	0.288***	0.314***	0.322***	0.159***
	(0.024)	(0.020)	(0.031)	(0.036)	(0.032)	(0.036)
$\log(M/A)$*ActHHI	**-0.026*** **	**-0.030*** **	**-0.027*** **	**-0.052*** **	**-0.059*** **	**-0.035*** **
	(0.005)	**(0.004)**	**(0.004)**	**(0.010)**	**(0.009)**	**(0.011)**
$\log(M/A)$*$PasHHI$	-0.036***	-0.034***	-0.036***	-0.035**	-0.028**	-0.029*
	(0.008)	(0.008)	(0.010)	(0.013)	(0.013)	(0.015)
$\log(M/A)$*IO	0.025***	0.019***	0.028***	0.027***	0.015*	0.041***
	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)	(0.009)
$\log(M/A)$*$Leverage$	-0.037***	-0.037***	-0.047***	-0.004	-0.013	-0.003
	(0.005)	(0.004)	(0.009)	(0.013)	(0.011)	(0.019)
$\log(M/A)$*$Sale$	0.021***	0.019***	0.019***	0.026***	0.023***	0.029***
	(0.003)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
$\log(M/A)$*$Cash$	-0.069***	-0.060***	-0.083***	-0.097***	-0.085***	-0.107***
	(0.012)	(0.008)	(0.011)	(0.012)	(0.009)	(0.014)
$ActHHI$	-0.012***	-0.010***	-0.022***	-0.013	-0.004	-0.023**
	(0.003)	(0.004)	(0.004)	(0.009)	(0.008)	(0.011)
$PasHHI$	0.030***	0.034***	0.015	0.076***	0.108***	0.010
	(0.009)	(0.009)	(0.011)	(0.019)	(0.018)	(0.021)
IO	-0.005	-0.007*	0.002	-0.028**	-0.041***	-0.012
	(0.004)	(0.004)	(0.006)	(0.011)	(0.011)	(0.009)
$Leverage$	0.054***	0.052***	0.064***	0.035**	0.025*	0.041**
	(0.005)	(0.005)	(0.010)	(0.014)	(0.013)	(0.017)
$Sale$	0.026***	0.024***	0.049***	0.049***	0.057***	0.046***
	(0.003)	(0.002)	(0.005)	(0.005)	(0.005)	(0.005)
$Cash$	0.011	-0.010	0.068***	0.061***	0.058***	0.067***
	(0.010)	(0.009)	(0.012)	(0.015)	(0.015)	(0.022)
Observations	83,054	83,794	83,952	69,612	70,250	70,402
R^2	0.823	0.837	0.714	0.677	0.697	0.579
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Table 3 – *Continued*

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel B: $ActTop5$			
$\log(M/A)$	0.056***	0.070***	0.059***	0.048***	0.084***	-0.004
	(0.012)	(0.007)	(0.014)	(0.016)	(0.013)	(0.023)
E/A	0.540***	0.560***	0.288***	0.322***	0.334***	0.162***
	(0.024)	(0.020)	(0.031)	(0.036)	(0.032)	(0.037)
$\log(M/A)$*$ActTop5$	**-0.033*** **	**-0.040*** **	**-0.029*** **	**-0.045*** **	**-0.063*** **	**-0.022***
	(0.006)	**(0.005)**	**(0.005)**	**(0.010)**	**(0.009)**	**(0.012)**
$\log(M/A)$*$PasTop5$	-0.016**	-0.013*	-0.013*	-0.025**	-0.005	-0.028**
	(0.006)	(0.006)	(0.007)	(0.012)	(0.011)	(0.013)
$\log(M/A)$*IO	0.021**	0.016**	0.027***	0.023***	0.014*	0.037***
	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)	(0.009)
$\log(M/A)$*$Leverage$	-0.035***	-0.035***	-0.046***	-0.001	-0.010	-0.001
	(0.005)	(0.004)	(0.009)	(0.013)	(0.011)	(0.019)
$\log(M/A)$*$Sale$	0.022***	0.019***	0.019***	0.027***	0.024***	0.029***
	(0.003)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
$\log(M/A)$*$Cash$	-0.070***	-0.062***	-0.084***	-0.097***	-0.087***	-0.105***
	(0.012)	(0.008)	(0.011)	(0.011)	(0.010)	(0.014)
$ActTop5$	-0.026***	-0.023***	-0.036***	-0.004	0.006	-0.023**
	(0.004)	(0.004)	(0.005)	(0.007)	(0.007)	(0.009)
$PasTop5$	0.045***	0.050***	0.026***	0.103***	0.138***	0.048***
	(0.006)	(0.006)	(0.006)	(0.013)	(0.015)	(0.012)
IO	-0.000	-0.001	0.003	-0.005	-0.011	0.000
	(0.004)	(0.004)	(0.006)	(0.010)	(0.010)	(0.009)
$Leverage$	0.052***	0.050***	0.062***	0.031**	0.019	0.039**
	(0.005)	(0.005)	(0.010)	(0.015)	(0.013)	(0.017)
$Sale$	0.025***	0.022***	0.049***	0.046***	0.053***	0.045***
	(0.003)	(0.002)	(0.005)	(0.005)	(0.005)	(0.005)
$Cash$	0.010	-0.011	0.067***	0.058***	0.054***	0.064***
	(0.010)	(0.009)	(0.012)	(0.014)	(0.015)	(0.022)
Observations	83,054	83,794	83,952	69,612	70,250	70,402
R^2	0.824	0.838	0.714	0.678	0.699	0.579
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Table 4: Investment-to-Price Sensitivity and Trading-based Ownership Concentration

This table reports OLS estimates on the relation between investment-to-price sensitivity, which measures the predictability of future investments from current market prices, and trading-based ownership concentration. The dependent variable is future investment rate, calculated as investment volume in year $t+h$ divided by total capital in year t. Here, h denotes the prediction horizons, set at 1 in Columns (1)-(3) and 3 in Columns (4)-(6). Investment volume is measured across one of the following three dimensions: (1) Intangible investment ($Intangible$) computed as R&D expense plus 30% SG&A expense; (2) Physical investment ($Physical$) captured by capital expenditure; (3) Total investment ($Invest$) representing the sum of $Physical$ and $Intangible$. Trading-based ownership concentration is measured by $ActHHI$ in Panel A, representing the Herfindahl-Hirschman index of active institutional ownership, and by $ActTop5$ in Panel B, denoting the proportion of shares held by the top five active institutional investors relative to the total shares held by all active institutional investors. $\log(M/A)$ is the log-ratio of a firm's market capitalization to its total assets. See Table B1 of the online appendix for the complete list of variable definitions. Standard errors, clustered at the year and firm levels, are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: $ActHHI$						
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
$\log(M/A)$	0.033***	0.054***	0.085***	0.074***	0.070***	0.139***
	(0.006)	(0.004)	(0.008)	(0.007)	(0.007)	(0.011)
I/K	1.091***	0.653***	0.927***	1.245***	0.378***	0.917***
	(0.049)	(0.025)	(0.039)	(0.084)	(0.041)	(0.062)
$\log(M/A)$*$ActHHI$	**-0.022*** **	**-0.023*** **	**-0.044*** **	**-0.042*** **	**-0.025*** **	**-0.070*** **
	(0.007)	**(0.003)**	**(0.007)**	**(0.009)**	**(0.004)**	**(0.010)**
$\log(M/A)$*$PasHHI$	0.013***	-0.001	0.016*	0.030**	-0.004	0.027
	(0.005)	(0.004)	(0.009)	(0.012)	(0.010)	(0.021)
$\log(M/A)$*IO	-0.022***	-0.028***	-0.050***	-0.054***	-0.038***	-0.094***
	(0.004)	(0.003)	(0.006)	(0.006)	(0.007)	(0.012)
$\log(M/A)$*$Leverage$	-0.016***	-0.006**	-0.024***	-0.043***	-0.005	-0.046***
	(0.003)	(0.003)	(0.004)	(0.008)	(0.006)	(0.013)
$\log(M/A)$*$Sale$	-0.003	-0.004***	-0.007**	0.003	-0.000	0.003
	(0.002)	(0.001)	(0.003)	(0.003)	(0.002)	(0.004)
$\log(M/A)$*$Cash$	0.072***	0.005	0.084***	0.136***	0.001	0.158***
	(0.013)	(0.006)	(0.020)	(0.013)	(0.009)	(0.022)
$ActHHI$	-0.014***	-0.015***	-0.025***	-0.001	0.006	0.014
	(0.003)	(0.003)	(0.005)	(0.007)	(0.005)	(0.010)
$PasHHI$	0.031***	0.030***	0.064***	0.142***	0.100***	0.256***
	(0.006)	(0.005)	(0.010)	(0.020)	(0.018)	(0.038)
IO	-0.010***	-0.013***	-0.025***	-0.063***	-0.059***	-0.137***
	(0.003)	(0.003)	(0.005)	(0.008)	(0.008)	(0.015)
$Leverage$	-0.008**	-0.015***	-0.026***	-0.031***	-0.036***	-0.071***
	(0.003)	(0.004)	(0.006)	(0.009)	(0.009)	(0.016)
$Sale$	-0.020***	-0.007***	-0.021***	-0.025***	0.004	-0.009
	(0.003)	(0.001)	(0.004)	(0.005)	(0.003)	(0.006)
$Cash$	0.018**	0.064***	0.100***	0.077***	0.100***	0.209***
	(0.008)	(0.007)	(0.013)	(0.017)	(0.013)	(0.029)
Observations	83,616	82,913	83,549	70,174	69,462	70,098
R^2	0.863	0.692	0.771	0.765	0.613	0.681
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Table 4 – *Continued*

	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: $ActTop5$						
$\log(M/A)$	0.040***	0.060***	0.096***	0.074***	0.067***	0.136***
	(0.010)	(0.006)	(0.014)	(0.011)	(0.008)	(0.015)
I/K	1.091***	0.655***	0.929***	1.242***	0.384***	0.921***
	(0.049)	(0.025)	(0.038)	(0.083)	(0.040)	(0.061)
$\log(M/A)$*$ActTop5$	**-0.027*** **	**-0.024*** **	**-0.050*** **	**-0.049*** **	**-0.028*** **	**-0.082*** **
	(0.009)	**(0.005)**	**(0.012)**	**(0.011)**	**(0.004)**	**(0.013)**
$\log(M/A)$*$PasTop5$	0.016***	0.011**	0.030***	0.050***	0.025***	0.081***
	(0.004)	(0.004)	(0.007)	(0.011)	(0.009)	(0.021)
$\log(M/A)$*IO	-0.022***	-0.027***	-0.048***	-0.045***	-0.028***	-0.074***
	(0.005)	(0.003)	(0.007)	(0.006)	(0.005)	(0.010)
$\log(M/A)$*$Leverage$	-0.017***	-0.007**	-0.025***	-0.045***	-0.006	-0.049***
	(0.003)	(0.003)	(0.004)	(0.008)	(0.006)	(0.012)
$\log(M/A)$*$Sale$	-0.003	-0.004***	-0.007**	0.002	-0.000	0.002
	(0.002)	(0.001)	(0.003)	(0.003)	(0.002)	(0.004)
$\log(M/A)$*$Cash$	0.071***	0.003	0.080***	0.130***	-0.004	0.147***
	(0.013)	(0.006)	(0.020)	(0.013)	(0.008)	(0.021)
$ActTop5$	-0.014***	-0.017***	-0.027***	-0.003	0.006	0.012
	(0.003)	(0.003)	(0.005)	(0.007)	(0.005)	(0.011)
$PasTop5$	0.029***	0.034***	0.067***	0.130***	0.114***	0.262***
	(0.004)	(0.005)	(0.008)	(0.014)	(0.015)	(0.030)
IO	-0.007**	-0.008***	-0.016***	-0.042***	-0.036***	-0.087***
	(0.003)	(0.003)	(0.005)	(0.007)	(0.006)	(0.012)
$Leverage$	-0.009***	-0.016***	-0.029***	-0.036***	-0.041***	-0.082***
	(0.003)	(0.004)	(0.007)	(0.009)	(0.009)	(0.017)
$Sale$	-0.021***	-0.007***	-0.022***	-0.027***	0.002	-0.013**
	(0.003)	(0.001)	(0.004)	(0.005)	(0.003)	(0.006)
$Cash$	0.018**	0.064***	0.099***	0.076***	0.098***	0.205***
	(0.008)	(0.007)	(0.013)	(0.016)	(0.013)	(0.028)
Observations	83,616	82,913	83,549	70,174	69,462	70,098
R^2	0.863	0.693	0.771	0.766	0.616	0.684
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Table 5: Identification based on Active Institutional Mergers

This table reports our identification results based on a quasi-natural experiment of financial institution mergers. Panel A validates our DiD model by testing the impact of financial institution mergers among active investors on two measures of trading-based ownership concentration: $ActHHI$ in Columns (1) and (3), and $ActTop5$ in Columns (2) and (4). $Treat$ is a treatment dummy, equal to 1 for firms held by both acquirer and target for more than 0.01% of the stock's market capitalization before the merger events. Control firms are those held by either the acquirer or the target, amounting to at least 0.01% of the market capitalization before the merger events. Besides, control firms are restricted to those that had never been treated in any of the merger events. $Post$ equals one for the post-merger period. The estimation is on a quarterly basis, with an estimation window of $(-8, +8)$ quarters in Columns (1)-(2) and $(-12, +12)$ quarters in Columns (3)-(4). Panel B (C) presents the relation between price informativeness (investment-to-price sensitivity) and trading-based ownership concentration using DiD models. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Panel A: Trading-based Ownership Concentration and Active Institutional Mergers

Event window	$(-8, +8)$ quarters		$(-12, +12)$ quarters	
Dependent variable	$ActHHI$	$ActTop5$	$ActHHI$	$ActTop5$
	(1)	(2)	(3)	(4)
$Treat$*$Post$	0.021***	0.034***	0.021***	0.034***
	(0.004)	(0.005)	(0.005)	(0.006)
Observations	95,396	95,396	135,524	135,524
R^2	0.605	0.690	0.536	0.633
Merger-Firm	Y	Y	Y	Y
Merger-Quarter FE	Y	Y	Y	Y

Panel B: DiD Estimation of Price Informativeness within $(-2, +2)$ years

where E =	$E_{h=1}/A$ $EBIT$	$E_{h=1}/A$ $EBITDA$	$E_{h=1}/A$ NI	$E_{h=3}/A$ $EBIT$	$E_{h=3}/A$ $EBITDA$	$E_{h=3}/A$ NI
	(1)	(2)	(3)	(4)	(5)	(6)
$\log(M/A)$*$Treat$*$Post$	-0.014***	-0.013***	-0.012***	-0.029***	-0.034***	-0.011
	(0.003)	(0.004)	(0.004)	(0.010)	(0.011)	(0.010)
Observations	23,504	23,661	23,737	21,563	21,660	21,738
R^2	0.839	0.846	0.737	0.741	0.757	0.624
Controls	Y	Y	Y	Y	Y	Y
Merger-Firm	Y	Y	Y	Y	Y	Y
Merger-Year FE	Y	Y	Y	Y	Y	Y

Panel C: DiD Estimation of Investment-to-Price Sensitivity within $(-2, +2)$ years

where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
	(1)	(2)	(3)	(4)	(5)	(6)
$\log(M/A)$*$Treat$*$Post$	-0.001	-0.010***	-0.010**	-0.023**	-0.016***	-0.042***
	(0.002)	(0.002)	(0.004)	(0.008)	(0.006)	(0.014)
Observations	23,619	23,360	23,607	21,638	21,317	21,623
R^2	0.897	0.748	0.812	0.847	0.725	0.770
Controls	Y	Y	Y	Y	Y	Y
Merger-Firm	Y	Y	Y	Y	Y	Y
Merger-Year FE	Y	Y	Y	Y	Y	Y

Table 6: International Evidence

This table utilizes the international sample to re-examine the relation between price informativeness and trading-based ownership concentration. The international sample is constructed by combining data on global institutional ownership from FactSet, accounting data from Worldscope, and stock market data from DataStream. Panels A and B focus on price informativeness, while Panels C and D focus on investment-to-price sensitivity. Trading-based ownership concentration is measured by $ActHHI$ in Panels A and C, representing the Herfindahl-Hirschman index of active institutional ownership, and by $ActTop5$ in Panels B and D, denoting the proportion of shares held by the top five active institutional investors relative to the total shares held by all active institutional investors. The sample possesses an annual frequency and spans from 2000 to 2022. The coefficients of the control variables are suppressed for brevity. Standard errors, clustered at the year and firm levels, are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Price Informativeness and $ActHHI$						
where E =	$\frac{E_{h=1}}{A}$ $EBIT$	$\frac{E_{h=1}}{A}$ $EBITDA$	$\frac{E_{h=1}}{A}$ NI	$\frac{E_{h=3}}{A}$ $EBIT$	$\frac{E_{h=3}}{A}$ $EBITDA$	$\frac{E_{h=3}}{A}$ NI
$\log(M/A)$*$ActHHI$	-0.021***	-0.021***	-0.018***	-0.025***	-0.026***	-0.022***
	(0.003)	(0.003)	(0.003)	(0.005)	(0.006)	(0.005)
Observations	172,863	172,514	178,447	141,518	141,203	146,716
R^2	0.725	0.733	0.701	0.617	0.630	0.598
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Country-Year	Y	Y	Y	Y	Y	Y
Panel B: Price Informativeness and $ActTop5$						
where E =	$\frac{E_{h=1}}{A}$ $EBIT$	$\frac{E_{h=1}}{A}$ $EBITDA$	$\frac{E_{h=1}}{A}$ NI	$\frac{E_{h=3}}{A}$ $EBIT$	$\frac{E_{h=3}}{A}$ $EBITDA$	$\frac{E_{h=3}}{A}$ NI
$\log(M/A)$*$ActTop5$	-0.029***	-0.029***	-0.024***	-0.037***	-0.041***	-0.031***
	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)	(0.005)
Observations	172,863	172,514	178,447	141,518	141,203	146,716
R^2	0.725	0.733	0.701	0.619	0.632	0.599
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Country-Year	Y	Y	Y	Y	Y	Y

Continued on next page

Table 6 – *Continued*

	(1)	(2)	(3)	(4)	(5)	(6)
Panel C: Investment-to-Price Sensitivity and $ActHHI$						
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
$\log(M/A)$*$ActHHI$	-0.007***	-0.015***	-0.022***	-0.022**	-0.023***	-0.049***
	(0.003)	(0.004)	(0.006)	(0.008)	(0.006)	(0.013)
Observations	178,293	177,640	177,992	146,622	146,034	146,391
R^2	0.846	0.623	0.679	0.715	0.545	0.581
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Country-Year	Y	Y	Y	Y	Y	Y
Panel D: Investment-to-Price Sensitivity and $ActTop5$						
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
$\log(M/A)$*$ActTop5$	-0.010***	-0.022***	-0.032***	-0.032***	-0.038***	-0.074***
	(0.003)	(0.004)	(0.005)	(0.008)	(0.008)	(0.016)
Observations	178,293	177,640	178,293	146,622	146,034	146,622
R^2	0.846	0.623	0.676	0.715	0.545	0.580
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Country-Year	Y	Y	Y	Y	Y	Y

Table 7: Portfolio Turnover of Active Institutional Investors

This table compares the portfolio turnover (PTR) in the Top5 subgroup and the Non-Top5 subgroup. The Top5 subgroup comprises stocks where the investor ranks among the top five largest shareholders, while the Non-Top5 subgroup includes all other stocks. Panel A illustrates the distribution of PTR for both the Top5 and Non-Top5 subgroups. The final two rows of Panel A adjust the threshold to be the top 10 ranking. Panel B presents regression analyses. Column (1) shows the results of regressing PTR on the dummy variable $DumTop5$, which is set to one for investor's Top5 subgroup, and zero for her Non-Top5 subgroup. Column (2) substitutes the Top5 subgroup dummy variable with the Top10 subgroup dummy variable, which is set to one for investor's Top10 subgroup. PIO is the portfolio institution ownership calculated as the holding-weighted average of stock-level institution ownership; $PRet$ is the portfolio quarterly return; $PRetStd$ is the portfolio volatility, calculated as the standard deviation of the quarterly returns in the past two years; $PSize$ is the portfolio size, computed as the logarithm of holding amount in million dollars. Standard errors, clustered at the quarter and investor levels, are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Panel A: Distribution of Portfolio Turnover

Subgroup	N	p10	p25	p50	p75	p90
Top5	69261	0.000	0.003	0.059	0.142	0.265
Non-Top5	69261	0.047	0.122	0.230	0.381	0.533
Top10	79249	0.000	0.020	0.095	0.190	0.326
Non-Top10	79249	0.039	0.116	0.231	0.396	0.560

Panel B: Regression of Portfolio Turnover

	(1) PTR	(2) PTR
$DumTop5$	-0.143***	
	(0.004)	
$DumTop10$		-0.125***
		(0.003)
PIO	0.046***	0.050***
	(0.010)	(0.010)
$PRet$	0.027***	0.037***
	(0.010)	(0.011)
$PRetStd$	-0.108***	-0.095***
	(0.019)	(0.022)
$PSize$	0.003**	0.001
	(0.001)	(0.001)
Observations	114,396	130,924
R^2	0.720	0.712
Investor-Quarter	Y	Y

Table 8: Information Content of Earnings Announcements

This table examines the relation between information content of earnings announcements and trading-based ownership concentration, as measured by $ActHHI$ and $ActTop5$. Information content is measured by abnormal trading volume ($AVOL$) in Columns (1)-(2) and abnormal return volatility ($AVAR$) in Columns (3)-(4). Specifically, $AVOL$ is calculated as the average trading volume in the event window $[0, 1]$, scaled by the counterparts in the non-event window $[-40, -6]$, where day 0 denotes the earnings announcement date; $AVAR$ is calculated as the mean square of adjusted returns in the event window, scaled by the counterparts in the non-event window. Panel B mirrors Panel A, with the addition of several control variables as specified by Pevzner, Xie, and Xin (2015) (abbreviated PXX): $FirmSize$ denotes the natural logarithm of the market capitalization at the fiscal quarter end; $|UE|$ is the absolute value of unexpected earnings, computed as actual annual earnings minus the most recent median analyst forecast scaled by the quarter-end stock price; $ReportLag$ is the number of days from the fiscal quarter-end to the earnings announcement date; $ForeDisp$ is the standard deviation of analysts' earnings forecasts scaled by the fiscal quarter-end stock price, and $ForeNum$ is the number of annual earnings forecasts reported by IBES. The coefficients of the control variables are suppressed for brevity. Standard errors, clustered at the quarter and firm levels, are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)
Panel A: Baseline Regression				
	$AVOL$	$AVOL$	$AVAR$	$AVAR$
$ActHHI$	-0.167***		-0.302***	
	(0.018)		(0.034)	
$ActTop5$		-0.169***		-0.301***
		(0.018)		(0.037)
Observations	319,619	319,619	320,050	320,050
R^2	0.263	0.263	0.245	0.246
Controls	Y	Y	Y	Y
PXX's Controls	N	N	N	N
Firm	Y	Y	Y	Y
Industry-Quarter	Y	Y	Y	Y
Panel B: Including PXX's Control Variables				
	$AVOL$	$AVOL$	$AVAR$	$AVAR$
$ActHHI$	-0.141***		-0.207***	
	(0.024)		(0.057)	
$ActTop5$		-0.101***		-0.125**
		(0.022)		(0.048)
Observations	162,570	162,570	162,575	162,575
R^2	0.313	0.313	0.271	0.271
Controls	Y	Y	Y	Y
PXX's Controls	Y	Y	Y	Y
Firm	Y	Y	Y	Y
Industry-Quarter	Y	Y	Y	Y

Internet Appendix to "Institutional Ownership Concentration and Informational Efficiency"

A Details of Numerical Analysis

This section details the calibration process for determining the ownership sizes λs of 21 investors across ten scenarios as mentioned at the end of Section 2.3.1. Investor 0 represents retail investors, while the remaining 20 investors are oligopolists. These oligopolists are further categorized into four distinct groups based on their size — large (L) or small (S) — and learning capacity — active (A) or passive (P). That is, $LA = \{1, 2\}$, $LP = \{3, 4\}$, $SA = \{5, \ldots, 12\}$, and $SP = \{13, \ldots, 20\}$.

We choose the mass of retail investors $\lambda_0 = 40\%$. The remaining 60% of ownership is allocated among the 20 oligopolists. Given that the sum of all ownership sizes must equal one, we have

$$\sum_{j \in LA} \lambda_j + \sum_{j \in SA} \lambda_j + \sum_{j \in LP} \lambda_j + \sum_{j \in SP} \lambda_j = 1 - \lambda_0 = 0.6. \tag{A1}$$

The relative size within each small group is set to be linearly distributed between 1 and 5. Specifically, $\frac{\lambda_5}{\lambda_{12}} = 5$ and $\frac{\lambda_{13}}{\lambda_{20}} = 5$. This allows us to express the total ownership of the small active and small passive investors as

$$\sum_{j \in SA} \lambda_j = 24\lambda_{12} \text{ and } \sum_{j \in SP} \lambda_j = 24\lambda_{20}, \tag{A2}$$

respectively. Furthermore, the passive sector is set at 50% of total institutional ownership:

$$\sum_{j \in LP} \lambda_j + \sum_{j \in SP} \lambda_j = 0.5(1 - \lambda_0) = 0.3. \tag{A3}$$

Next, we hold the total size of active section ($\sum_{j \in LA} \lambda_j + \sum_{j \in SA} \lambda_j$) and the total size of passive section ($\sum_{j \in LP} \lambda_j + \sum_{j \in SP} \lambda_j$) fixed, and reshuffle the ownership sizes λs within each section create varying levels of ownership concentration. Specifically, we vary two sets of parameter values. First, we change the relative size of the two large active oligopolists and two large passive oligopolists by varying $a_{act} \equiv \lambda_1/\lambda_2$ and $a_{pas} \equiv \lambda_3/\lambda_4$ linearly from 1.1 to 10 in ten scenarios. That is, $a_{act}(k) = a_{pas}(k) = 1.1 + \frac{10-1.1}{9}(k-1)$, where k denotes the scenario. Consequently, we can express the total ownership of the large active and passive investors, respectively,

as

$$\sum_{j \in LA} \lambda_j = (1 + a_{act}(k))\lambda_2 \text{ and } \sum_{j \in LP} \lambda_j = (1 + a_{pas}(k))\lambda_4. \tag{A4}$$

Second and at the same time, we vary the relative size of the small sector linearly from 10% to 3% in the ten scenarios. More formally:

$$\sum_{j \in SA} \lambda_j = (1 - 0.5)(1 - \lambda_0) b_{act}(k) = 0.3 b_{act}(k), \tag{A5}$$

$$\sum_{j \in SP} \lambda_j = 0.5(1 - \lambda_0) b_{pas}(k) = 0.3 b_{pas}(k), \tag{A6}$$

where $b_{act}(k) = b_{pas}(k) = 0.10 - \frac{0.10-0.03}{9}(k-1)$.

Finally, with equations (A1)–(A6), we can solve for λ_j for each investor in each scenario k. Table A1 reports the ownership sizes λ_j along with the size-based ownership concentration values for the ten scenarios. Passive sized-based ownership concentration is calculated as $PasHHI_{size} = \sum_{j \in SP \cup LP} \left(\frac{\lambda_j}{\sum_{k \in SP \cup LP} \lambda_k} \right)^2$.

While alternative methods exist to generate varying concentration levels, our simultaneous adjustment of parameters a and b is motivated by observed patterns in our empirical sample. Specifically, we categorize institutional investors in our sample into 20 groups, mirroring the 20 oligopolists in our calibration experiments. This allows us to examine the evolution of their size distribution over time.

Regarding categorization, we use active investors as an example to illustrate. We begin by ranking active investors annually based on their AUM. The top 5% of active investors are classified as "Large." Within this group, the top 1% are assigned to Group 1, while the remaining investors are placed in Group 2. The remaining 95% of active investors are considered "Small." These small investors are further divided into 8 equally sized groups (Groups 5 through 12) based on their AUM. We then calculate the average AUM for each group, denoted as AUM_j for investor $j = \{1, 2, 5, \ldots, 12\}$.[1] We repeat this procedure for passive investors, creating an analogous set

[1] Instead of using AUM deciles, we define large investors as those in the top 5% by AUM. This better captures the role of the exceptionally large investors. Our results are robust to alternative thresholds, such as the top 3% and top 10%.

of groups.

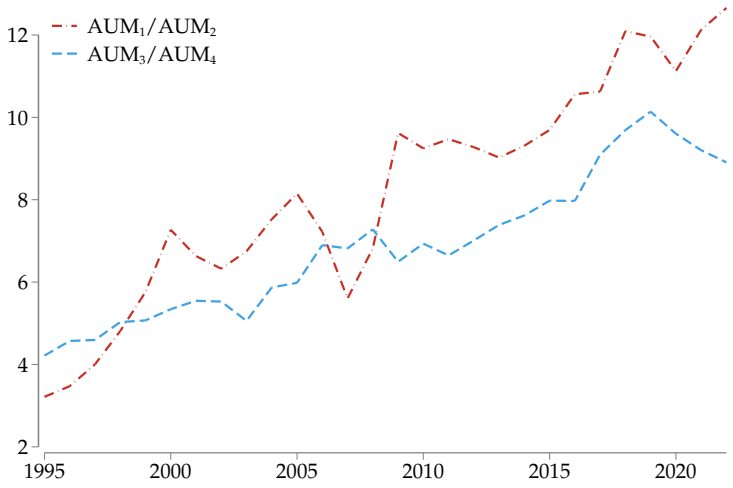
Figure A1 presents the time-series patterns of the relative sizes within the large sectors ($\frac{AUM_1}{AUM_2}$ and $\frac{AUM_3}{AUM_4}$) and the relative size of the small sectors ($\frac{AUM_{j\in SA}}{AUM_{j\in SA\cup LA}}$ and $\frac{AUM_{j\in SP}}{AUM_{j\in SP\cup LP}}$). We observe an overall increase in the relative size within the large sectors and a concurrent decrease in the relative size of the small sectors over time. These trends motivate our method of simultaneously varying a and b to generate increasing concentration in the calibration experiments.

We then compare the distribution of λ_j across the 10 calibration scenarios with the evolution of AUM_j in our empirical sample. Panels (a) and (c) of Figure A2 illustrate the evolution of AUM_j within the active and passive sectors, respectively.[2] Notably, these distributions closely resemble the distribution of λ_j across the 10 calibration scenarios, as shown in Panels (c) and (d). Therefore, the simultaneous adjustment of parameters a and b makes the size distribution in the calibration exercise match well with our empirical sample.

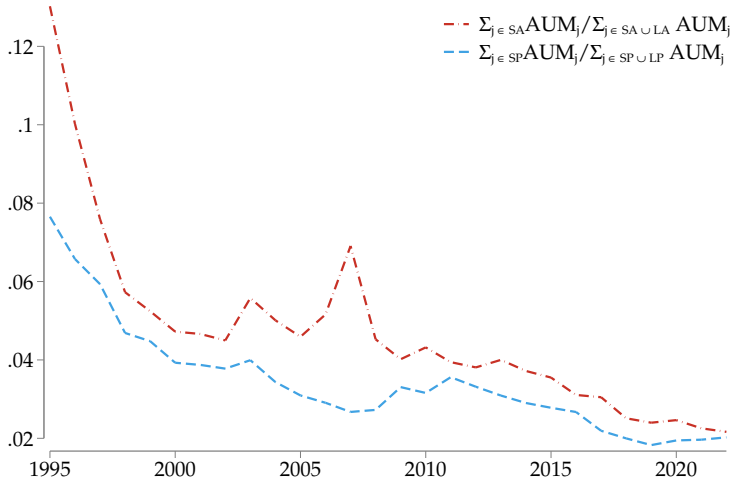
In the placebo tests mentioned in Section 2.3.4, we respectively fix the size distribution for active and passive investors throughout ten scenarios. For the first placebo test, the relative size of the two large active investors is fixed at 1.1, i.e., $a_{act}(k) = 1.1$ for $k = \{1, 2, \ldots, 10\}$. Meanwhile, the relative size of the small active sector is fixed at 0.10, i.e., $b_{act}(k) = 0.10$ for $k = \{1, 2, \ldots, 10\}$. We then resolve equations (A1)–(A6) to obtain λ_j in the first placebo test. Table A2 reports the results. While $PasHHI_{size}$ increases gradually from 0.4074 to 0.7855 across the ten scenarios, $ActHHI_{size}$ remains fixed at 0.4047.

For the second placebo test, the relative size of the two large passive investors is fixed at 1.1, i.e., $a_{pas}(k) = 1.1$ for $k = \{1, 2, \ldots, 10\}$. Meanwhile, the relative size of the small passive sector is fixed at 0.10, i.e., $b_{pas}(k) = 0.10$ for $k = \{1, 2, \ldots, 10\}$. We then resolve equations (A1)–(A6) again to obtain λ_j in the second placebo test. Table A3 reports the results. While $ActHHI_{size}$ increases gradually from 0.4074 to 0.7855 across the ten scenarios, $PasHHI_{size}$ remains fixed at 0.4047.

[2]The sample begins in 1995 due to the relatively small number of active investors prior to 1995 (less than 300), which would result in limited observations within each group.

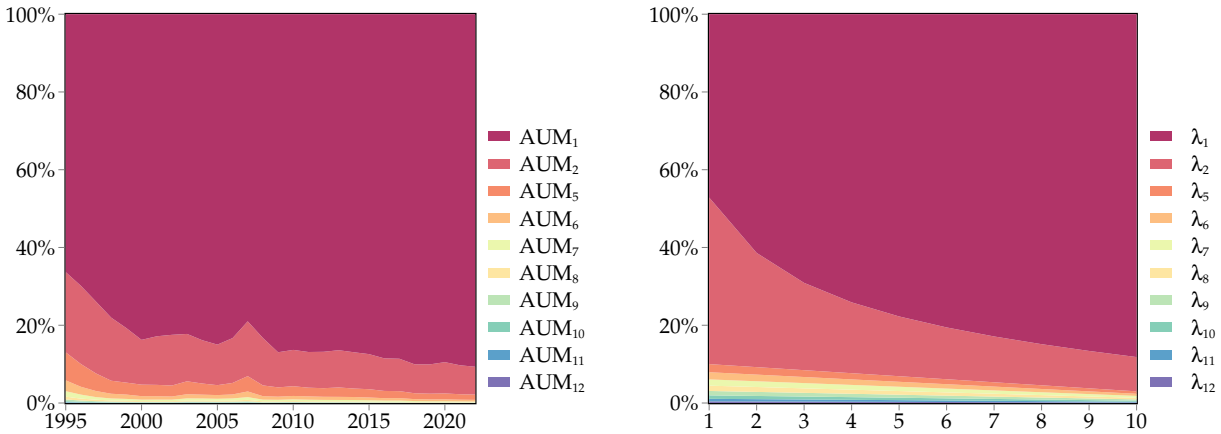


(a) Relative Size within the Large Sectors



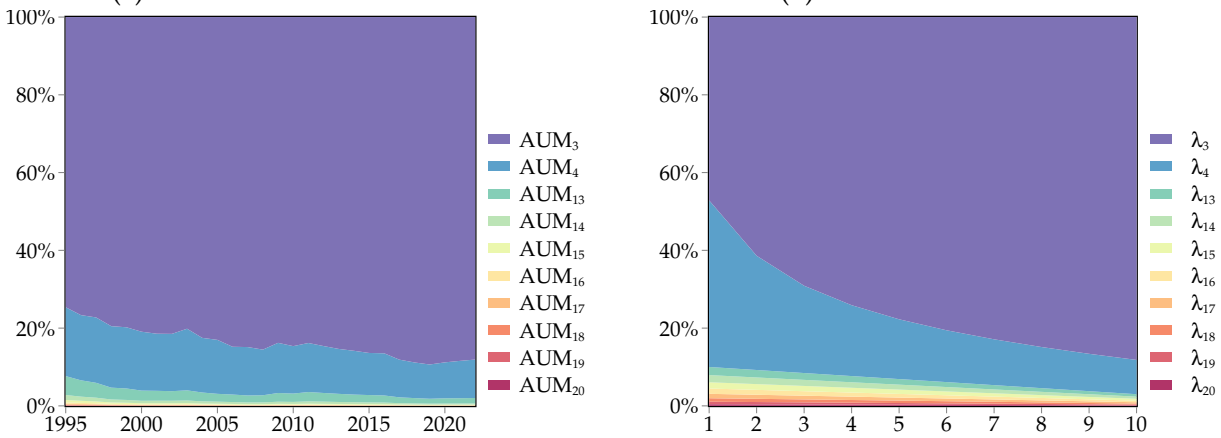
(b) Relative Size of the Small Sectors

Figure A1: Relative Size Distribution



(a) AUM of Active Investors

(b) λ of Active Investors



(c) AUM of Passive Investors

(d) λ of Passive Investors

Figure A2: Size Distribution

Table A1: Ownership Sizes and Size-based Concentration in Ten Scenarios: Baseline

Investor Type		1	2	3	4	5	6	7	8	9	10
Panel A: λ_j											
Retail	λ_0	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000
Large Active	λ_1	0.1414	0.1842	0.2073	0.2223	0.2332	0.2417	0.2486	0.2546	0.2598	0.2645
	λ_2	0.1286	0.0882	0.0674	0.0547	0.0461	0.0400	0.0354	0.0317	0.0288	0.0265
Large Passive	λ_3	0.1414	0.1842	0.2073	0.2223	0.2332	0.2417	0.2486	0.2546	0.2598	0.2645
	λ_4	0.1286	0.0882	0.0674	0.0547	0.0461	0.0400	0.0354	0.0317	0.0288	0.0265
Small Active	λ_5	0.0063	0.0058	0.0053	0.0048	0.0043	0.0038	0.0033	0.0028	0.0024	0.0019
	λ_6	0.0055	0.0051	0.0047	0.0042	0.0038	0.0034	0.0030	0.0025	0.0021	0.0017
	λ_7	0.0048	0.0044	0.0041	0.0037	0.0033	0.0029	0.0026	0.0022	0.0018	0.0014
	λ_8	0.0041	0.0038	0.0035	0.0031	0.0028	0.0025	0.0022	0.0019	0.0016	0.0012
	λ_9	0.0034	0.0031	0.0029	0.0026	0.0023	0.0021	0.0018	0.0015	0.0013	0.0010
	λ_{10}	0.0027	0.0025	0.0023	0.0021	0.0018	0.0016	0.0014	0.0012	0.0010	0.0008
	λ_{11}	0.0020	0.0018	0.0017	0.0015	0.0014	0.0012	0.0010	0.0009	0.0007	0.0006
	λ_{12}	0.0013	0.0012	0.0011	0.0010	0.0009	0.0008	0.0007	0.0006	0.0005	0.0004
Small Passive	λ_{13}	0.0063	0.0058	0.0053	0.0048	0.0043	0.0038	0.0033	0.0028	0.0024	0.0019
	λ_{14}	0.0055	0.0051	0.0047	0.0042	0.0038	0.0034	0.0030	0.0025	0.0021	0.0017
	λ_{15}	0.0048	0.0044	0.0041	0.0037	0.0033	0.0029	0.0026	0.0022	0.0018	0.0014
	λ_{16}	0.0041	0.0038	0.0035	0.0031	0.0028	0.0025	0.0022	0.0019	0.0016	0.0012
	λ_{17}	0.0034	0.0031	0.0029	0.0026	0.0023	0.0021	0.0018	0.0015	0.0013	0.0010
	λ_{18}	0.0027	0.0025	0.0023	0.0021	0.0018	0.0016	0.0014	0.0012	0.0010	0.0008
	λ_{19}	0.0020	0.0018	0.0017	0.0015	0.0014	0.0012	0.0010	0.0009	0.0007	0.0006
	λ_{20}	0.0013	0.0012	0.0011	0.0010	0.0009	0.0008	0.0007	0.0006	0.0005	0.0004
Panel B: $ActHHI_{size}$ and $PasHHI_{size}$											
$ActHHI_{size}$		0.4074	0.4645	0.5290	0.5833	0.6286	0.6673	0.7013	0.7317	0.7596	0.7855
$PasHHI_{size}$		0.4074	0.4645	0.5290	0.5833	0.6286	0.6673	0.7013	0.7317	0.7596	0.7855

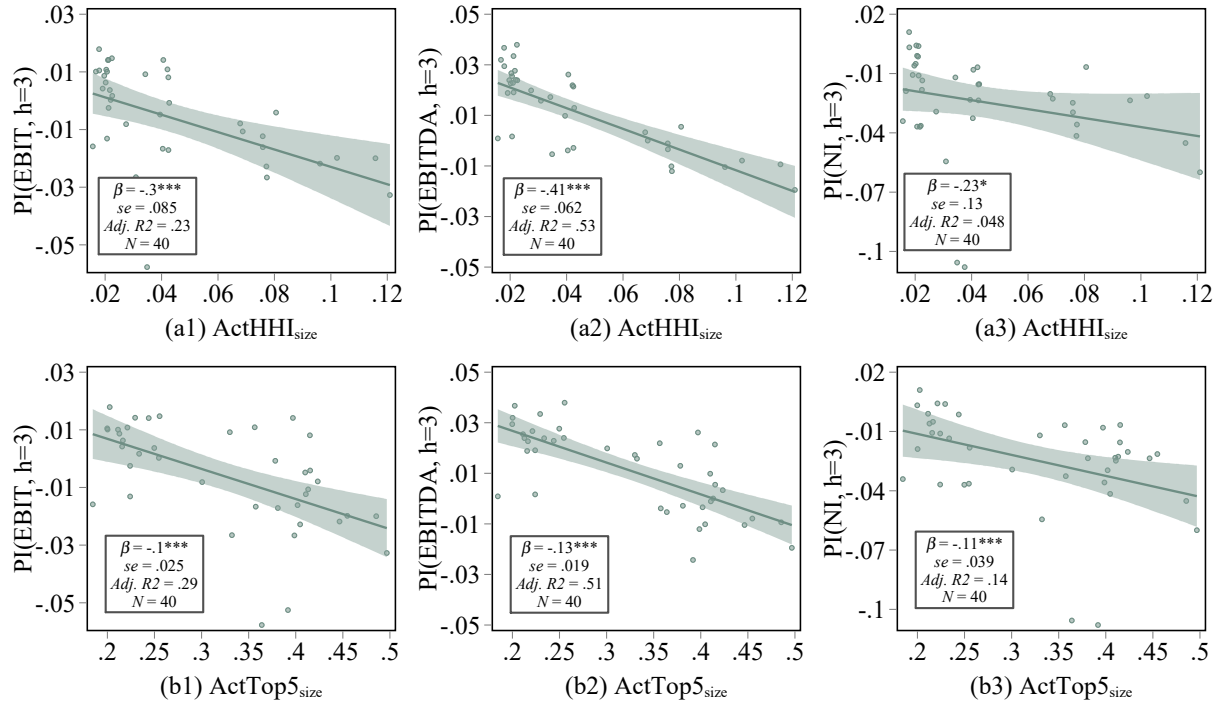
Table A2: Ownership Sizes and Size-based Concentration: Fixed $ActHHI_{size}$

Investor Type		1	2	3	4	5	6	7	8	9	10
Panel A: λ_j											
Retail	λ_0	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000
Large Active	λ_1	0.1414	0.1414	0.1414	0.1414	0.1414	0.1414	0.1414	0.1414	0.1414	0.1414
	λ_2	0.1286	0.1286	0.1286	0.1286	0.1286	0.1286	0.1286	0.1286	0.1286	0.1286
Large Passive	λ_3	0.1414	0.1842	0.2073	0.2223	0.2332	0.2417	0.2486	0.2546	0.2598	0.2645
	λ_4	0.1286	0.0882	0.0674	0.0547	0.0461	0.0400	0.0354	0.0317	0.0288	0.0265
Small Active	λ_5	0.0063	0.0063	0.0063	0.0063	0.0063	0.0063	0.0063	0.0063	0.0063	0.0063
	λ_6	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055
	λ_7	0.0048	0.0048	0.0048	0.0048	0.0048	0.0048	0.0048	0.0048	0.0048	0.0048
	λ_8	0.0041	0.0041	0.0041	0.0041	0.0041	0.0041	0.0041	0.0041	0.0041	0.0041
	λ_9	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034
	λ_{10}	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027
	λ_{11}	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020
	λ_{12}	0.0013	0.0013	0.0013	0.0013	0.0013	0.0013	0.0013	0.0013	0.0013	0.0013
Small Passive	λ_{13}	0.0063	0.0058	0.0053	0.0048	0.0043	0.0038	0.0033	0.0028	0.0024	0.0019
	λ_{14}	0.0055	0.0051	0.0047	0.0042	0.0038	0.0034	0.0030	0.0025	0.0021	0.0017
	λ_{15}	0.0048	0.0044	0.0041	0.0037	0.0033	0.0029	0.0026	0.0022	0.0018	0.0014
	λ_{16}	0.0041	0.0038	0.0035	0.0031	0.0028	0.0025	0.0022	0.0019	0.0016	0.0012
	λ_{17}	0.0034	0.0031	0.0029	0.0026	0.0023	0.0021	0.0018	0.0015	0.0013	0.0010
	λ_{18}	0.0027	0.0025	0.0023	0.0021	0.0018	0.0016	0.0014	0.0012	0.0010	0.0008
	λ_{19}	0.0020	0.0018	0.0017	0.0015	0.0014	0.0012	0.0010	0.0009	0.0007	0.0006
	λ_{20}	0.0013	0.0012	0.0011	0.0010	0.0009	0.0008	0.0007	0.0006	0.0005	0.0004
Panel B: $ActHHI_{size}$ and $PasHHI_{size}$											
$ActHHI_{size}$		0.4074	0.4074	0.4074	0.4074	0.4074	0.4074	0.4074	0.4074	0.4074	0.4074
$PasHHI_{size}$		0.4074	0.4645	0.5290	0.5833	0.6286	0.6673	0.7013	0.7317	0.7596	0.7855

Table A3: Ownership Sizes and Size-based Concentration in Ten Scenarios: Fixed $PasHHI_{size}$

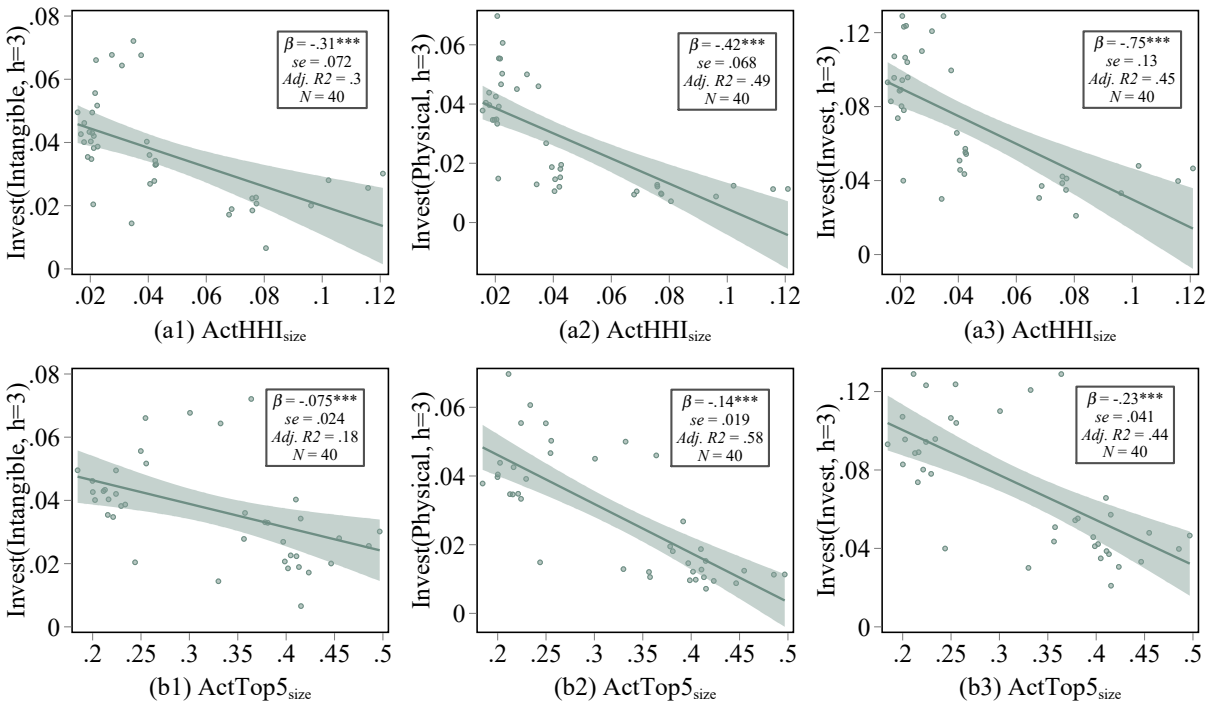
Investor Type		1	2	3	4	5	6	7	8	9	10
Panel A: λ_j											
Retail	λ_0	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000
Large Active	λ_1	0.1414	0.1842	0.2073	0.2223	0.2332	0.2417	0.2486	0.2546	0.2598	0.2645
	λ_2	0.1286	0.0882	0.0674	0.0547	0.0461	0.0400	0.0354	0.0317	0.0288	0.0265
Large Passive	λ_3	0.1414	0.1414	0.1414	0.1414	0.1414	0.1414	0.1414	0.1414	0.1414	0.1414
	λ_4	0.1286	0.1286	0.1286	0.1286	0.1286	0.1286	0.1286	0.1286	0.1286	0.1286
Small Active	λ_5	0.0063	0.0058	0.0053	0.0048	0.0043	0.0038	0.0033	0.0028	0.0024	0.0019
	λ_6	0.0055	0.0051	0.0047	0.0042	0.0038	0.0034	0.0030	0.0025	0.0021	0.0017
	λ_7	0.0048	0.0044	0.0041	0.0037	0.0033	0.0029	0.0026	0.0022	0.0018	0.0014
	λ_8	0.0041	0.0038	0.0035	0.0031	0.0028	0.0025	0.0022	0.0019	0.0016	0.0012
	λ_9	0.0034	0.0031	0.0029	0.0026	0.0023	0.0021	0.0018	0.0015	0.0013	0.0010
	λ_{10}	0.0027	0.0025	0.0023	0.0021	0.0018	0.0016	0.0014	0.0012	0.0010	0.0008
	λ_{11}	0.0020	0.0018	0.0017	0.0015	0.0014	0.0012	0.0010	0.0009	0.0007	0.0006
	λ_{12}	0.0013	0.0012	0.0011	0.0010	0.0009	0.0008	0.0007	0.0006	0.0005	0.0004
Small Passive	λ_{13}	0.0063	0.0063	0.0063	0.0063	0.0063	0.0063	0.0063	0.0063	0.0063	0.0063
	λ_{14}	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055	0.0055
	λ_{15}	0.0048	0.0048	0.0048	0.0048	0.0048	0.0048	0.0048	0.0048	0.0048	0.0048
	λ_{16}	0.0041	0.0041	0.0041	0.0041	0.0041	0.0041	0.0041	0.0041	0.0041	0.0041
	λ_{17}	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034
	λ_{18}	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027
	λ_{19}	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020
	λ_{20}	0.0013	0.0013	0.0013	0.0013	0.0013	0.0013	0.0013	0.0013	0.0013	0.0013
Panel B: $ActHHI_{size}$ and $PasHHI_{size}$											
$ActHHI_{size}$		0.4074	0.4645	0.5290	0.5833	0.6286	0.6673	0.7013	0.7317	0.7596	0.7855
$PasHHI_{size}$		0.4074	0.4074	0.4074	0.4074	0.4074	0.4074	0.4074	0.4074	0.4074	0.4074

B Additional Figures and Tables in Main Text



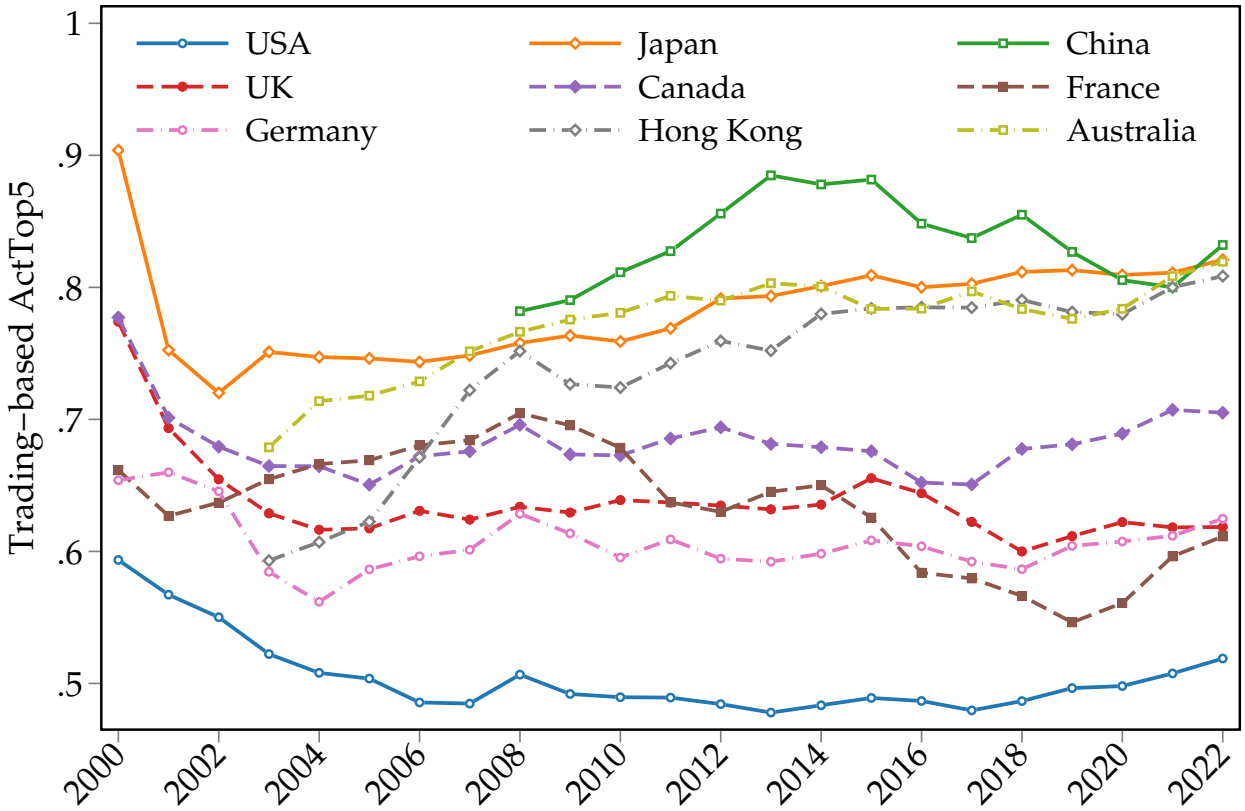
This figure presents scatter plots illustrating the relation between price informativeness and size-based ownership concentration. The plots include fit lines and 95% confidence intervals. Size-based ownership concentration is quantified using $ActHHI_{size}$ in Panels (a1)-(a3), and $ActTop5_{size}$ in Panels (b1)-(b3). PI is derived from equations (12) and (13) and measures the predictability of future cash flows based on current market prices, with future cash flows represented by one of three variables ($EBIT$, $EBITDA$, or NI) calculated as of year $t+h$ and divided by total assets in year t. The prediction horizon, denoted by h, is set at 3 years. The sample has an annual frequency and spans from 1980 to 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Figure B1: Price Informativeness and Size-based Ownership Concentration: 3-year Prediction Horizon



This figure presents scatter plots illustrating the relation between investment-to-price sensitivity and size-based ownership concentration. The plots include fit lines and 95% confidence intervals. Size-based ownership concentration is quantified using $ActHHI_{size}$ in Panels (a1)-(a3), and $ActTop5_{size}$ in Panels (b1)-(b3). Investment-to-price sensitivity is derived from equation (14) and measures the extent to which current market prices reveal the information necessary for future investment decisions, with future investments represented by one of three variables ($Intangible$, $Physical$, or $Invest$) calculated as of year $t + h$ and divided by total capital in year t. The prediction horizon, denoted by h, is set at 3 years. The sample has an annual frequency and spans from 1980 to 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Figure B2: Investment-to-Price Sensitivity and Size-based Ownership Concentration: 3-year Prediction Horizon



This figure displays the time-series average trading-based ownership concentration ($ActTop5$) for the largest equity markets worldwide, including the United States, the United Kingdom, Germany, Japan, Canada, Hong Kong, China, France, and Australia.

Figure B3: The Time Trend of Trading-based Ownership Concentration Worldwide

Table B1: Variable Definitions

Variable	Description
$ActHHI_{size}$	Herfindahl-Hirschman index of Assets Under Management (AUM) among active institutional investors: $$ActHHI_{size,q} = \frac{\sum_{j=1}^{N_{mkt}} \left(AUM_{j,q}^2\right)}{\left(\sum_{j=1}^{N} AUM_{j,q}\right)^2},$$ where N is the total number of institutional investors; $AUM_{j,q}$ is the AUM of institution j in quarter q. The definition of active and passive institutional investors is based on the classfication scheme of Bushee (1998).
$ActTop5_{size}$	The proportion of AUM held by the top five active institutional investors relative to the total AUM of all active institutional investors: $$ActTop5_{size,q} = \frac{\sum_{j=1}^{\text{Top 5}} AUM_{j,q}}{\sum_{j=1}^{N_{mkt}} AUM_{j,q}},$$ where N is the total number of institutional investors; $AUM_{j,q}$ is the AUM of institution j in quarter q.
$ActHHI$	Herfindahl-Hirschman index of active institutional ownership: $$ActHHI_{i,q} = \frac{\sum_{j=1}^{N_i} \left(S_{i,j,q}^2\right)}{(\sum_{j=1}^{N_i} S_{i,j,q})^2},$$ where $S_{i,j,q}$ denotes the equity shares of stock i owned by active institution j in quarter q; N_i is the number of active institutions holding stock i.
$ActTop5$	The proportion of shares held by the top five active institutional investors relative to the total shares held by all active institutional investors.: $$ActTop5_{i,q} = \frac{\sum_{j=1}^{\text{Top 5}} S_{i,j,q}}{\sum_{j=1}^{N_i} S_{i,j,q}},$$ where $S_{i,j,q}$ denotes the equity shares of stock i owned by active institution j in quarter q; N_i is the number of active institutions holding stock i.
$EBIT/A$	Earnings before interest and taxes scaled by total assets.
$EBITDA/A$	Earnings before interest, taxes, depreciation and amortization scaled by total assets.
NI/A	Net income scaled by total assets.

Continued on next page

Table B1 – *Continued*

Variable	Description
$Intangible/K$	Intangible investment rate, calculated as R&D + 0.3 $\times$ SG&A expenses, scaled by total capital. R&D is set to zero for missing values. The total capital is the sum of net property, plant and equipment (item PPENT from Compustat) and intangible capital (item K_INT from Peters and Taylor (2017)).
$Physical/K$	Physical investment rate, calculated as capital expenditure (CAPX) scaled by total capital. The total capital is the sum of net property, plant and equipment (item PPENT from Compustat) and intangible capital (item K_INT from Peters and Taylor (2017)).
$Invest/K$	Total investment rate, defined as the sum of $Physical$ and $Intangible$.
$\log(M/A)$	The log-ratio of market capitalization at the end of March to the total asset value in the previous fiscal year.
$PasHHI$	Herfindahl-Hirschman index of passive institutional ownership. The calculation method closely resembles that of $ActHHI$, with the key distinction being the transition from active to passive investors within the cohort considered.
$PasTop5$	Holding percentage of the largest five passive shareholders. The calculation method closely resembles that of $ActTop5$, with the key distinction being the transition from active to passive investors within the cohort considered.
IO	Institutional ownership, calculated as the total institution holding divided by the market capitalization.
$Leverage$	Ratio of book debt to total assets.
$Sale$	Total sales scaled by total assets.
$Cash$	Cash holdings scaled by total assets.
$BHAR[\tau, T]$	Buy-and-hold abnormal returns from day τ to day T ($\tau < T$), where day 0 denotes the earnings announcement day.
$Rank$	A decile rank of the analyst earnings surprises, with analyst earnings surprises calculated as the difference between the quarter's actual earnings per share and the median of the latest analyst forecasts, divided by the firm's stock price five trading days prior to the announcement date.
$CPIE$	A microstructure-based measure developed by Duarte et al. (2020), capturing the probability of private information arrival on a given day. The measure is derived from one of the four microstructure models of private information arrival: the PIN model (PIN) of Easley et al. (1996), the adjusted PIN model (APIN) of Duarte and Young (2009), the generalized PIN model (GPIN) of Duarte et al. (2020), and the Odders-White and Ready (2008) model (OWR).

Continued on next page

Table B1 – *Continued*

Variable	Description
ITI	A machine learning-based measure of informed trading intensity by Bogousslavsky et al. (2024). The measure is trained from one of the three samples: Schedule 13D trading, opportunistic insider trades, and short sales.
$VR(q)$	A q-period bias-corrected variance ratio by Lo and MacKinlay (1988): $$VR(q) = \left\lvert \frac{\sigma^2(q)}{q \times \sigma^2} - 1 \right\rvert,$$ where $\sigma^2(q)$ denotes the variance of returns over a q-day horizon; σ^2 denotes the variance of daily returns.
$\tau_{\pi}^{R,j}$	A measure of relative price informativeness by Dávila and Parlatore (2025), which corresponds to the Kalman gain of a Bayesian external observer who only learns from the price under a Gaussian environment.
PTR	Portfolio turnover, calculated as $$PTR_{k,g,q} = \frac{\min\left(AgBuy_{k,g,q}, AgSell_{k,g,q}\right)}{\sum_{i \in N_{k,g}} (S_{k,g,i,q} P_{i,q} + S_{k,g,i,q-1} P_{i,q-1})/2},$$ where $AgBuy_{k,g,q}$ and $AgSell_{k,g,q}$ are the aggregate purchase and sale of portfolio g held by active institutional investor k in quarter q, respectively; S is the number of holding shares; P is the share price.
$AVAR$	Abnormal return volatility, calculated as the mean of the squared market-model-adjusted returns in the event window (earnings announcement event), scaled by the counterparts in the non-event window.
$AVOL$	Abnormal trading volume, calculated as the mean of share turnover in the event window (earnings announcement event), scaled by the counterparts in the non-event window.

Table B2: Active Financial Institution Mergers

This table lists the sample of 11 active financial institution mergers that are used for identification, including the announcement date, completion date, acquirer name and target name of the merger.

Announcement Date	Completion Date	Acquirer Name	Target Name
1986/7/15	1986/8/28	Travelers Corp	Dillon Read & Co Inc
1995/5/8	1995/12/27	U.S. Bancorp	West One Bank, Idaho NA
1996/4/15	1996/4/30	Equitable Life Assurance	Natl Mutual Funds Mgmt
1996/6/24	1996/10/31	Morgan Stanley Group Inc	Van Kampen Amer Capital
1997/11/5	1997/12/1	PIMCO Advisors LP	Oppenheimer Group Inc
2003/7/22	2003/10/31	Lehman Brothers Hldgs	Neuberger Berman, LLC (Sloate)
2003/10/14	2004/2/27	Hennessy Advr Inc	Lindner Asset Management, Inc
2004/8/26	2005/1/31	Blackrock Inc	State Str Research & Mgmt Co
2010/2/16	2010/4/19	Fortress Invt Grp, LLC	Guggenheim Capital, LLC
2017/5/9	2017/10/2	Two Sigma Secs, LLC	Timber Hill LLC
2018/4/10	2018/4/10	Schonfeld Strategic Advr LLC	Folger Hill Asset Mgmt LLC

Table B3: DiD Estimation Using Active Institutional Mergers: Alternative Control Group

This table replicates our DiD results from Table 5, with an alternative strategy for selecting control firms: Control firms are re-defined as those held by the acquirer but not the target, with a 0.01% or greater ownership prior to the merger announcement. Panel A focuses on price informativeness, while Panel B focuses on investment-to-price sensitivity. $Treat$ is a treatment dummy, equal to 1 for firms held by both acquirer and target for more than 0.01% of the stock's market capitalization before the merger events. $Post$ equals one for the post-merger period. The estimation is conducted on an annual basis, with an estimation window from 2 years before to 2 years after mergers. The coefficients of the control variables are suppressed for brevity. See Table B1 for the complete list of variable definitions. Standard errors, clustered at the year and firm levels, are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Price Informativeness within $(-2, +2)$ years						
where E =	$\frac{E_{h=1}}{A}$ $EBIT$	$\frac{E_{h=1}}{A}$ $EBITDA$	$\frac{E_{h=1}}{A}$ NI	$\frac{E_{h=3}}{A}$ $EBIT$	$\frac{E_{h=3}}{A}$ $EBITDA$	$\frac{E_{h=3}}{A}$ NI
$\log(M/A)$*$Treat$*$Post$	-0.016***	-0.015***	-0.013***	-0.031***	-0.035***	-0.014
	(0.003)	(0.004)	(0.004)	(0.010)	(0.011)	(0.010)
Observations	20,740	20,897	20,967	19,021	19,118	19,190
R^2	0.840	0.848	0.739	0.750	0.765	0.633
Controls	Y	Y	Y	Y	Y	Y
Merger-Firm	Y	Y	Y	Y	Y	Y
Merger-Year FE	Y	Y	Y	Y	Y	Y
Panel B: Investment-to-Price Sensitivity within $(-2, +2)$ years						
where I=	$\frac{I_{h=1}}{K}$ $Intangible$	$\frac{I_{h=1}}{K}$ $Physical$	$\frac{I_{h=1}}{K}$ $Invest$	$\frac{I_{h=3}}{K}$ $Intangible$	$\frac{I_{h=3}}{K}$ $Physical$	$\frac{I_{h=3}}{K}$ $Invest$
$\log(M/A)$*$Treat$*$Post$	-0.000	-0.010***	-0.009**	-0.020**	-0.017***	-0.039***
	(0.002)	(0.002)	(0.004)	(0.009)	(0.006)	(0.014)
Observations	20,849	20,600	20,840	19,090	18,784	19,078
R^2	0.897	0.746	0.810	0.845	0.723	0.770
Controls	Y	Y	Y	Y	Y	Y
Merger-Firm	Y	Y	Y	Y	Y	Y
Merger-Year FE	Y	Y	Y	Y	Y	Y

Table B4: DiD Estimation Using Active Institutional Mergers: Alternative Event Window

This table replicates our DiD results from Table 5, with the key distinction that the estimation window is extended to (-3, +3) years, with year-0 denoting the merger completion year. Panel A focuses on price informativeness, while Panel B focuses on investment-to-price sensitivity. $Treat$ is a treatment dummy, equal to 1 for firms held by both acquirer and target for more than 0.01% of the stock's market capitalization before the merger events. Control firms are those held by either the acquirer or the target, amounting to at least 0.01% of the market capitalization before the merger events. Besides, control firms are restricted to those that had never been treated in any of the merger events. $Post$ equals one for the post-merger period. The estimation is conducted on an annual basis, with an estimation window from 3 years before to 3 years after mergers. The coefficients of the control variables are suppressed for brevity. See Table B1 for the complete list of variable definitions. Standard errors, clustered at the year and firm levels, are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Price Informativeness within $(-3, +3)$ years						
where E =	$E_{h=1}/A$ $EBIT$	$E_{h=1}/A$ $EBITDA$	$E_{h=1}/A$ NI	$E_{h=3}/A$ $EBIT$	$E_{h=3}/A$ $EBITDA$	$E_{h=3}/A$ NI
$\log(M/A)$*$Treat$*$Post$	-0.012**	-0.012***	-0.009**	-0.038***	-0.042***	-0.019*
	(0.005)	(0.004)	(0.004)	(0.010)	(0.011)	(0.010)
Observations	32,001	32,210	32,310	29,416	29,557	29,658
R^2	0.810	0.820	0.688	0.700	0.713	0.585
Controls	Y	Y	Y	Y	Y	Y
Merger-Firm	Y	Y	Y	Y	Y	Y
Merger-Year	Y	Y	Y	Y	Y	Y
Panel B: Investment-to-Price Sensitivity within $(-3, +3)$ years						
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
$\log(M/A)$*$Treat$*$Post$	0.000	-0.008**	-0.007	-0.025**	-0.022***	-0.046***
	(0.002)	(0.004)	(0.005)	(0.010)	(0.007)	(0.017)
Observations	32,156	31,776	32,139	29,527	29,097	29,505
R^2	0.876	0.712	0.789	0.803	0.658	0.716
Controls	Y	Y	Y	Y	Y	Y
Merger-Firm	Y	Y	Y	Y	Y	Y
Merger-Year FE	Y	Y	Y	Y	Y	Y

Table B5: Summary Statistics for the International Sample

This table presents the summary statistics for the variables in the international sample. The international sample is constructed by combining data on global institutional ownership from FactSet, accounting data from Worldscope, and stock market data from DataStream. The sample has an annual frequency and spans from 2000 to 2022. All continuous variables are winsorized at the top and bottom 1% to mitigate the influence of outliers. Variable definitions are provided in Table B1.

Variable	N	Mean	SD	p10	p25	p50	p75	p90
Panel A: Ownership Concentration Variables								
$ActHHI$	196120	0.202	0.189	0.036	0.063	0.137	0.277	0.468
$ActTop5$	196120	0.675	0.241	0.333	0.462	0.691	0.908	0.988
Panel B: Earning Variables								
$EBIT/A$	191854	0.050	0.158	-0.060	0.027	0.067	0.114	0.176
$EBITDA/A$	191574	0.091	0.156	-0.014	0.060	0.105	0.158	0.224
NI/A	196112	0.019	0.155	-0.078	0.009	0.040	0.078	0.128
Panel C: Investment Rate Variables								
$Intangible/K$	195916	0.143	0.132	0.007	0.046	0.119	0.195	0.293
$Physical/K$	195533	0.100	0.106	0.014	0.033	0.068	0.128	0.221
$Invest/K$	195755	0.245	0.154	0.097	0.151	0.212	0.294	0.420
Panel D: Control Variables								
$\log(M/A)$	196120	-0.101	0.996	-1.342	-0.751	-0.110	0.555	1.186
$PasHHI$	196120	0.833	0.165	0.589	0.705	0.868	1.000	1.000
$PasTop5$	196120	0.366	0.293	0.098	0.147	0.243	0.508	0.966
IO	196120	0.319	0.322	0.028	0.069	0.177	0.485	0.928
$Leverage$	196120	0.218	0.188	0.000	0.047	0.193	0.339	0.475
$Sale$	196120	0.939	0.654	0.264	0.491	0.806	1.213	1.763
$Cash$	196120	0.189	0.186	0.021	0.056	0.129	0.256	0.452

Table B6: International Evidence: Exclude U.S. Firms

This table utilizes the international sample excluding firms in the United States to re-examine the relation between price informativeness and trading-based ownership concentration. The international sample is constructed by combining data on global institutional ownership from FactSet, accounting data from Worldscope, and stock market data from DataStream. Panels A and B focus on price informativeness, while Panels C and D focus on investment-to-price sensitivity. Trading-based ownership concentration is measured by $ActHHI$ in Panels A and C, representing the Herfindahl-Hirschman index of active institutional ownership, and by $ActTop5$ in Panels B and D, denoting the proportion of shares held by the top five active institutional investors relative to the total shares held by all active institutional investors. The sample possesses an annual frequency and spans from 2000 to 2022. The coefficients of the control variables are suppressed for brevity. See Table B1 for the complete list of variable definitions. Standard errors, clustered at the year and firm levels, are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Price Informativeness and $ActHHI$						
where E =	$E_{h=1}/A$ $EBIT$	$E_{h=1}/A$ $EBITDA$	$E_{h=1}/A$ NI	$E_{h=3}/A$ $EBIT$	$E_{h=3}/A$ $EBITDA$	$E_{h=3}/A$ NI
$\log(M/A)$*$ActHHI$	-0.020***	-0.020***	-0.014***	-0.020***	-0.021***	-0.015***
	(0.003)	(0.003)	(0.002)	(0.005)	(0.006)	(0.005)
Observations	120,376	120,170	123,243	99,368	99,160	102,132
R^2	0.694	0.708	0.666	0.599	0.625	0.572
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Country-Year FE	Y	Y	Y	Y	Y	Y
Panel B: Price Informativeness and $ActTop5$						
where E =	$E_{h=1}/A$ $EBIT$	$E_{h=1}/A$ $EBITDA$	$E_{h=1}/A$ NI	$E_{h=3}/A$ $EBIT$	$E_{h=3}/A$ $EBITDA$	$E_{h=3}/A$ NI
$\log(M/A)$*$ActTop5$	-0.030***	-0.030***	-0.022***	-0.042***	-0.046***	-0.034***
	(0.004)	(0.004)	(0.003)	(0.006)	(0.007)	(0.005)
Observations	120,376	120,170	123,243	99,368	99,160	102,132
R^2	0.695	0.709	0.666	0.602	0.628	0.574
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Country-Year FE	Y	Y	Y	Y	Y	Y

Continued on next page

Table B6 – *Continued*

	(1)	(2)	(3)	(4)	(5)	(6)
Panel C: Investment-to-Price Sensitivity and $ActHHI$						
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
$\log(M/A)$*$ActHHI$	-0.005***	-0.011**	-0.017***	-0.015**	-0.015*	-0.031**
	(0.002)	(0.004)	(0.005)	(0.007)	(0.008)	(0.014)
Observations	123,162	122,660	122,906	102,078	101,598	101,867
R^2	0.822	0.627	0.652	0.686	0.553	0.568
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Country-Year FE	Y	Y	Y	Y	Y	Y
Panel D: Investment-to-Price Sensitivity and $ActTop5$						
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
$\log(M/A)$*$ActTop5$	-0.008***	-0.019***	-0.027***	-0.023***	-0.033***	-0.055***
	(0.002)	(0.005)	(0.005)	(0.007)	(0.009)	(0.015)
Observations	123,162	122,660	122,906	102,078	101,598	101,867
R^2	0.822	0.626	0.652	0.686	0.553	0.568
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Country-Year FE	Y	Y	Y	Y	Y	Y

C Robustness Checks

In this section, we conduct several tests to demonstrate the robustness of the negative effect of trading-based ownership concentration on price informativeness.

C.1 Alternative Measures of Price Informativeness

Our baseline analysis follows Bai et al. (2016) to measure price informativeness. Although this particular measure is closely related to our theoretical analysis and has a strong economic appeal as a welfare measure under Q-theory, there is no general consensus on how to measure price informativeness. Therefore, we explore several alternative measures of price informativeness and demonstrate the robustness of the negative impact of trading-based ownership concentration.

C.1.1 Post-Earnings-Announcement Drift (PEAD)

To attenuate the concern of model misspecification, we consider post-earnings-announcement drift (PEAD), a model-free measure of price informativeness. Our sample of earnings announcement starts in 1984 due to the data availability of analyst forecast in I/B/E/S, and ends in 2022. We construct scaled earnings surprises following Akey et al. (2022):

$$SUE_{i,t} = \frac{EPS_{i,t} - E_{t-1}\left[EPS_{i,t}\right]}{P_{i,t-5}}, \tag{C1}$$

where $EPS_{i,t}$ is the earnings per share for firm i announced on day t, and $E_{t-1}[EPS_{i,t}]$ is the expectation of earnings per share, measured by the median of all analyst forecasts issued over the 90 days before the earnings announcement date.[3] If analysts revise their forecasts during this interval, only their most recent forecasts are included. We scale the surprise by the firm's stock price five trading days before the announcement.

[3]We collect earnings announcement dates from Compustat and I/B/E/S and go through the following steps to determine the effective date on which earnings announcements are made. First, we compare the announcement dates in the two databases and pick up the earlier one. Second, we eliminate cases where the earning announcement dates in the two databases are more than two trading days apart. Third, if the earnings are released prior to 4:00 PM Eastern Time from Monday through Friday according to the time stamp in I/B/E/S, the corresponding date is designated as the effective announcement date. Conversely, if the earnings are released at or after 4:00 PM Eastern Time from Monday through Friday, over the weekend, or on a trading holiday, the next trading date in CRSP is designated as the effective announcement date.

To quantify the efficiency of stock prices in incorporating earnings surprises on the announcement date, we first construct buy-and-hold abnormal returns for firm i's earnings announcement from day τ to day T ($\tau < T$) as $BHAR[\tau, T] = \prod_{k=\tau}^{T}(1+R_{i,k}) - \prod_{k=\tau}^{T}(1+R_{p,k})$, where the daily stock return $R_{i,k}$ is adjusted by the return on the size and book-to-market matching Fama-French portfolio $R_{p,k}$. Specifically, stocks are matched to one of 25 portfolios every year based on their market capitalization and book-to-market ratio. Market capitalization is calculated at the end of June, whereas the book-to-market ratio is calculated as the book equity of the last fiscal year end in the prior calendar year divided by the market value of equity at the end of December of the previous year.

Martineau (2022) shows that stock prices have become more efficient in incorporating earnings surprises in the last decade, especially for large stocks, as $BHAR$ jumped on the announcement date and has remained essentially flat for the following sixty trading days. We take a further step to study the interaction effect of ownership concentration on price efficiency by estimating the following regression models:

$$\begin{aligned} BHAR[0,2]_{i,t} =& \beta_1 Rank_{i,t} + \beta_2 Rank_{i,t} \times Concentration_{i,t} + \beta_3 Concentration_{i,t} + \\ & \rho \chi_{i,t} + FE_{i,t} + \varepsilon_{i,t}, \end{aligned} \tag{C2}$$

$$\begin{aligned} BHAR[3,24]_{i,t} =& \gamma_1 Rank_{i,t} + \gamma_2 Rank_{i,t} \times Concentration_{i,t} + \gamma_3 Concentration_{i,t} + \\ & \rho \chi_{i,t} + FE_{i,t} + \varepsilon_{i,t}, \end{aligned} \tag{C3}$$

where $BHAR[0,2]_{i,t}$ and $BHAR[3,24]_{i,t}$ correspond to firm i's announcement date and post-announcement BHAR, respectively. $Rank_{i,t}$ is a decile rank of the analyst earnings surprises defined in equation (C1). Decile ranks are established for each year-quarter by utilizing observations from the preceding quarter to define the decile breakpoints, thereby mitigating any potential look-ahead bias. As claimed by Martineau (2022), the decile rank is preferred compared to the original earnings surprise, because the distribution of earnings surprises has high kurtosis relative to a normal- or t-distribution.

Our coefficients of interest are β_2 in equation (C2) and γ_2 in equation (C3). If ownership

concentration impedes the efficiency of stock prices in incorporating earnings surprises around the announcement date, β_2 is expected to be negative. At the same time, we would expect a more persistent price drift as indicated by a positive γ_2. Panel A of Table C1 presents the results. Consistent with our hypothesis, stocks with higher concentration of active institutional ownership exhibit a smaller response of BHAR to earnings surprises around the announcement, along with larger price drifts. The result holds for two different measures of ownership concentration, namely, $ActHHI$ and $ActTop5$.

C.1.2 Conditional Probability of An Information Event (CPIE)

We consider a microstructure-based price informativeness measure developed by Duarte et al. (2020), $CPIE$, which captures the probability of private information arrival on a given day, conditional on the estimated structural model parameters and the observed daily stock characteristics. Specifically, the authors consider four microstructure models of private information arrival: the PIN model (PIN) of Easley et al. (1996), the adjusted PIN model (APIN) of Duarte and Young (2009), the generalized PIN model (GPIN) of Duarte et al. (2020), and the Odders-White and Ready (2008) model (OWR).[4] The authors estimate each of these models for each stock per year to obtain the structural parameters, and then calculate the daily $CPIE$ as the probability of an information event given the estimated structural parameters, as well as the observed daily order flows and stock returns for each stock.

We aggregate $CPIE$ to the stock-quarter level by taking the average, and regress it on the ownership concentration at the end of each quarter. Owing to the data availability of CPIE, our sample commences on January 4, 1993, and concludes on December 31, 2012. Panel B of Table C1 reports the results. From Columns (1) to (4), $CPIE$ is calculated based on the PIN, APIN, GPIN, and OWR models, respectively. Consistent with our hypothesis, the coefficients on $ActHHI$ are negative and statistically significant at the 1% level except for Column (4), suggesting that active

[4]The PIN model identifies private information based on order flow imbalance. The APIN model is a mixture of two independent PIN models, which allows the intensity of noise-trade arrivals to vary. In contrast to the APIN model, the GPIN model allows the noise trade intensity to vary continuously. While the PIN, APIN, and GPIN model only rely on order flow to infer whether private information has arrived, the OWR model takes into account the intra-day and overnight returns as well. See Duarte et al. (2020) for a more detailed discussion.

institutional ownership concentration lowers the probability of informed trading. The results are robust if we use $ActTop5$ as an alternative measure of the active institutional ownership concentration, as shown in Columns (5)-(8) in Panel B of Table C1.

C.1.3 Informed Trading Intensity

We also consider a machine learning-based measure of informed trading intensity (ITI) developed by Bogousslavsky et al. (2024). The authors define informed trading days as those that involve Schedule 13D trading, significant opportunistic insider trading, and significant short selling. They use a Gradient Boosted Trees (GBT) algorithm incorporating 41 concurrent daily variables (related to liquidity, return, volatility, and volume) to detect informed trading days. The developed ITI measure increases before earnings, M&A, and news announcements, indicating its effectiveness in detecting informed trading.

We collect firm-level daily ITI indexes from the authors' website and aggregate them to the firm-quarterly level by simply taking the average. Due to the data availability of ITI indexes, our sample period is from January 5, 1993 to July 31, 2019. We regress ITI indexes on the active institutional ownership concentration at the end of each quarter. Panel C of Table C1 reports the results. From Columns (1) to (3), the ITI measure is trained on informed trading samples of Schedule 13D trades, opportunistic insiders, and short sellers, respectively. The coefficients on $ActHHI$ remain significantly negative across all specifications, suggesting that stocks with more concentrated active institutional ownership are associated with less informed trading activities. The result remains robust if we use $ActTop5$ as the alternative measure of ownership concentration.

C.1.4 Variance Ratio

We next consider a weak-form price efficiency measure. Under perfect weak-form efficiency, stock prices evolve according to a random walk. A testable prediction of the random walk hypothesis is that returns over a q-day horizon should have a variance ($\sigma^2(q)$) that is q times the variance of daily returns (σ^2). Formally, we use the q-period bias-corrected variance ratio of Lo

and MacKinlay (1988):

$$VR(q) = \left| \frac{\sigma^2(q)}{q \times \sigma^2} - 1 \right|, \tag{C4}$$

where $\sigma^2 = \frac{1}{nq-1} \sum_{k=1}^{nq} (X_k - X_{k-1} - \hat{\mu})^2$, $\sigma^2(q) = \frac{1}{m} \sum_{k=q}^{nq} (X_k - X_{k-q} - q\hat{\mu})^2$, $\hat{\mu} = \frac{1}{nq} \sum_{k=1}^{nq} (X_k - X_{k-1})$, and $m = q(nq - q + 1)(1 - \frac{q}{nq})$. X denotes the log price, n denotes the number of nonoverlapping q-period returns in the measurement interval, and nq denotes the number of daily returns in the measurement interval. When prices follow a random walk, $VR(q)$ equals 0. The higher the value of $VR(q)$, the further the stock price process deviates from a random walk. If ownership concentration undermines weak-form price efficiency, we should obtain a positive relation between $VR(q)$ in equation (C4) and ownership concentration.

We conduct our tests using stock-quarter-level observations. More specifically, we first compute variance ratios over horizons of q = 5, 10, 15, and 20 trading days using overlapping observations during a quarter. We then regress them on the active institutional ownership concentration controlling for the firm fixed effect and industry-quarter fixed effects. Panel D of Table C1 shows that the variance ratio increases with active institutional ownership concentration, consistent with a lower price efficiency for stocks with more concentrated active institutional ownership. The results hold for different estimation horizons and are statistically significant at the 1% level.

C.1.5 Relative Price Informativeness

Dávila and Parlatore (2025) identify a measure of relative price informativeness, which corresponds to the Kalman gain of a Bayesian external observer who only learns from the price under a Gaussian environment. Formally, the authors employ the following panel regression models:

$$\Delta p_t^j = \bar{\beta}\left(Y_t^j\right) + \beta_0\left(Y_t^j\right) \Delta x_t^j + \beta_1\left(Y_t^j\right) \Delta x_{t+4}^j + \varepsilon_t^j,$$
$$\Delta p_t^j = \bar{\zeta}\left(Y_t^j\right) + \zeta_0\left(Y_t^j\right) \Delta x_t^j + \hat{\varepsilon}_t^j,$$

where Δp_t^j is the year-on-year change in log-price of stock j in quarter t; Δx_t^j and its one-year ahead counterpart Δx_{t+4}^j are measures of earnings growth, calculated as the logarithm of one plus

the year-on-year changes in EBIT divided by book equity; The coefficients are modeled as affine functions of firm-specific characteristics Y_t^j. The error variances specific to each firm $\mathbb{V}\mathrm{ar}[\varepsilon_t^j]$ and $\mathbb{V}\mathrm{ar}[\hat{\varepsilon}_t^j]$ are estimated respectively using $\widehat{\mathbb{V}\mathrm{ar}}\left[\varepsilon_t^j\right] = \exp\left\{\lambda_0 + \lambda_1 Y_t^j + Y_t^{j\prime}\lambda_2 Y_t^j + \frac{1}{2}\mathbb{V}\mathrm{ar}[v_t^j]\right\}$ and $\widehat{\mathbb{V}\mathrm{ar}}\left[\hat{\varepsilon}_t^j\right] = \exp\left\{\hat{\lambda}_0 + \hat{\lambda}_1 Y_t^j + Y_t^{j\prime}\hat{\lambda}_2 Y_t^j + \frac{1}{2}\mathbb{V}\mathrm{ar}[\hat{v}_t^j]\right\}$. Finally, the relative price informativeness for stock j in quarter t is quantified by $\hat{\tau}_{\pi,t}^{R,j} = \frac{\widehat{\mathbb{V}\mathrm{ar}}\left[\hat{\varepsilon}_t^j\right] - \widehat{\mathbb{V}\mathrm{ar}}\left[\varepsilon_t^j\right]}{\widehat{\mathbb{V}\mathrm{ar}}\left[\hat{\varepsilon}_t^j\right]}$.

The sample selection procedure is similar to that in our baseline analysis, expect for the additional requirement that stocks' relative price informativeness should be positive. Following Dávila and Parlatore (2025), we conduct our tests at the portfolio level. Specifically, we divide the sample into twenty bins each quarter based on the ownership concentration of each firm, and then aggregate the quarterly measures of relative price informativeness within each bin-quater. We conduct panel regressions of relative price informativeness on the ownership concentration variables at the bin-quarter level, controlling for the quarter fixed effect. The results in Panel E of Table C1 echo those in Table 2 in Dávila and Parlatore (2025). The coefficients on $ActHHI$ and $ActTop5$ are significantly negative, indicating that portfolios with more concentrated ownership have lower relative price informativeness. To control for the size effect, we take the residual from the regression of relative price informativeness on size with quarter fixed effect before running the panel regressions. As shown in the last two rows in Panel E of Table C1, the results remain significantly negative. Figures C1 and C2 provide alternative graphical illustrations of our results, indicating that the cross-sectional relations identified in Panel E of Table C1 are stable over time.

C.2 Alternative Sample: Mutual Fund Holdings

Form 13F filings are filed at the management company level rather than at the portfolio or individual fund level (Agarwal et al., 2013). This poses a challenge as a fund management company may oversee both passive and active mutual funds, potentially leading to measurement errors in the classification method proposed by Bushee (1998). To address this issue, we utilize fund-level holdings data from Thomson Reuters S12 as an alternative source to distinguish between active and passive mutual funds. While the S12 data provide a more precise measure of active/passive

ownership, it does not encompass other institutional investors beyond mutual fund management companies, such as banks, insurance companies, pension funds, and independent investment advisors. Thus, we rely on 13F holdings data for our primary analysis, using S12 data as a supplementary check for robustness.

Following previous studies (e.g., Appel et al., 2016), we flag a fund as passively managed if its fund name includes a string that identifies it as an index fund or if the CRSP Mutual Fund Database classifies the fund as an index fund.[5] Table C2 replicates the baseline results by using S12 holdings data. The coefficients on the interaction term are negative and statistically significant throughout, suggesting that our result is robust to the alternative definition of institutional investor at the disaggregated level.

C.3 Other Robustness Tests

Table C3 replicates our baseline results in Tables 3-4, with the distinction that we use Bushee's time-varying classification scheme to distinguish active/passive institutional investors, which updates the classification for every year in our sample period. The results are virtually unchanged.

While we compute trading-based ownership concentration based on the detailed holding data of each institution in our baseline analysis, Table C4 shows that our results remain robust if we calculate it based on each institution's trading volume in each firm's stock, as shown equation (8) defined in our model.

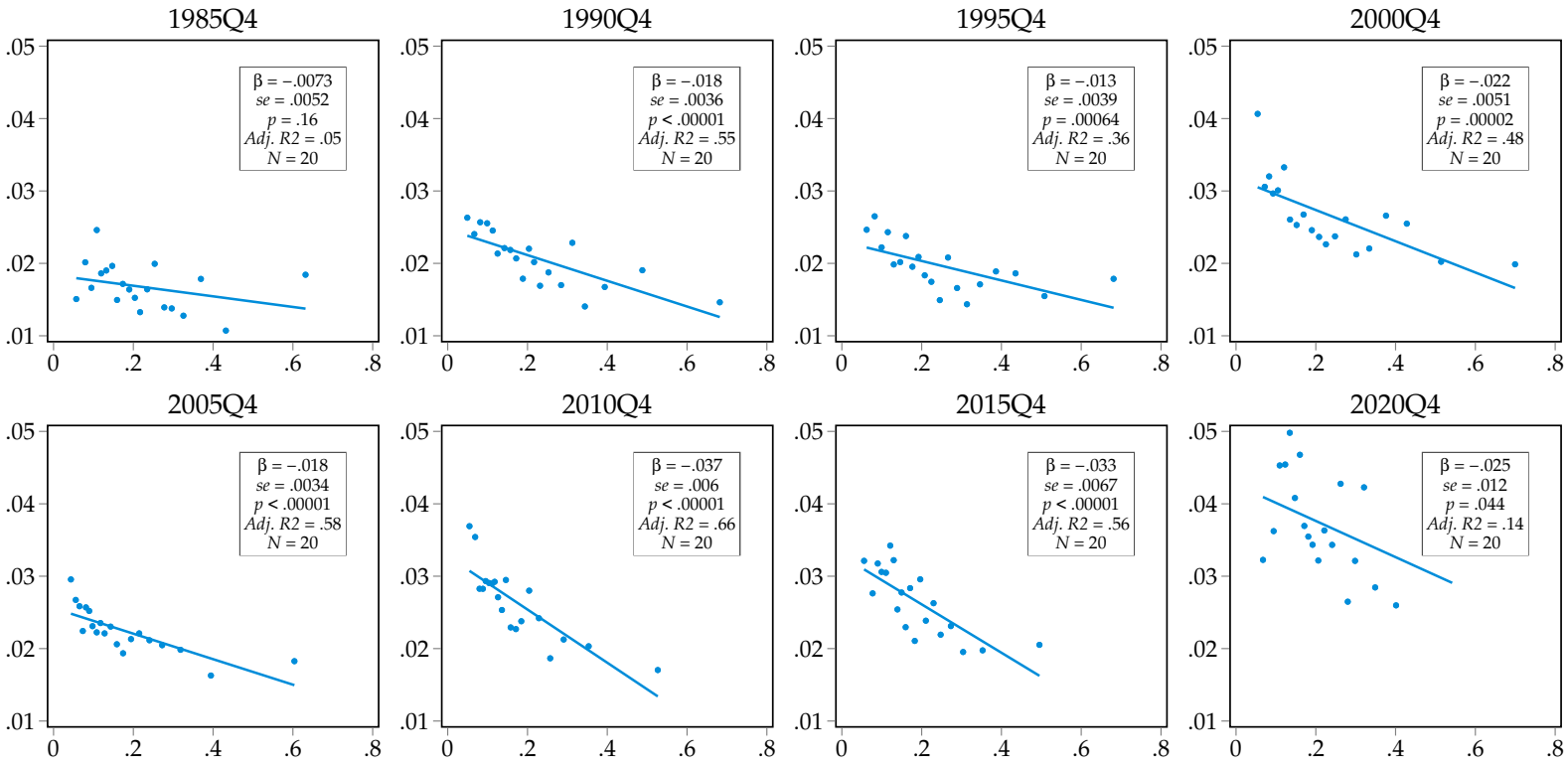
Although passive institutional investors do not directly affect the information level of stock prices as indicated in equation (6), they may indirectly affect price informativeness through the trading activities of active investors, as discussed in Section 2.3. Our baseline analysis accounts for this potential effect by controlling for the passive institutional ownership ($PasHHI$ or $PasTop5$). Alternatively, Table C5 reconstructs our concentration measures without distin-

[5]The strings we use to identify index funds include: Index, Idx, Indx, Ind_ (where _ indicates a space), Russell, S & P, S and P, S&P, SandP, SP, DOW, Dow, DJ, MSCI, Bloomberg, KBW, NASDAQ, NYSE, STOXX, FTSE, Wilshire, Morningstar, 100, 400, 500, 600, 900, 1000, 1500, 2000, and 5000. In addition, in CRSP, a fund with flag D is a "pure index fund" whose "objective is to match the total investment performance of a publicly recognized securities market index."

guishing between active and passive investors. Specifically, $TotHHI_{i,q} = \frac{\sum_{j=1}^{N_{tot}} \left(S_{i,j,q}^2\right)}{(\sum_{j=1}^{N_{tot}} S_{i,j,q})^2}$ captures HHI of institutional shares, where N_{tot} denotes the number of institutions holding stock i. Similarly, $TotTop5_{i,q} = \frac{\sum_{j=1}^{\text{Top 5}} S_{i,j,q}}{\sum_{j=1}^{N_{tot}} S_{i,j,q}}$ measures the proportion of shares held by the top five largest institutional investors relative to the total shares held by all institutional investors. We continue to observe a significant negative effect of trading-based ownership concentration on both price informativeness and investment-to-price sensitivity.

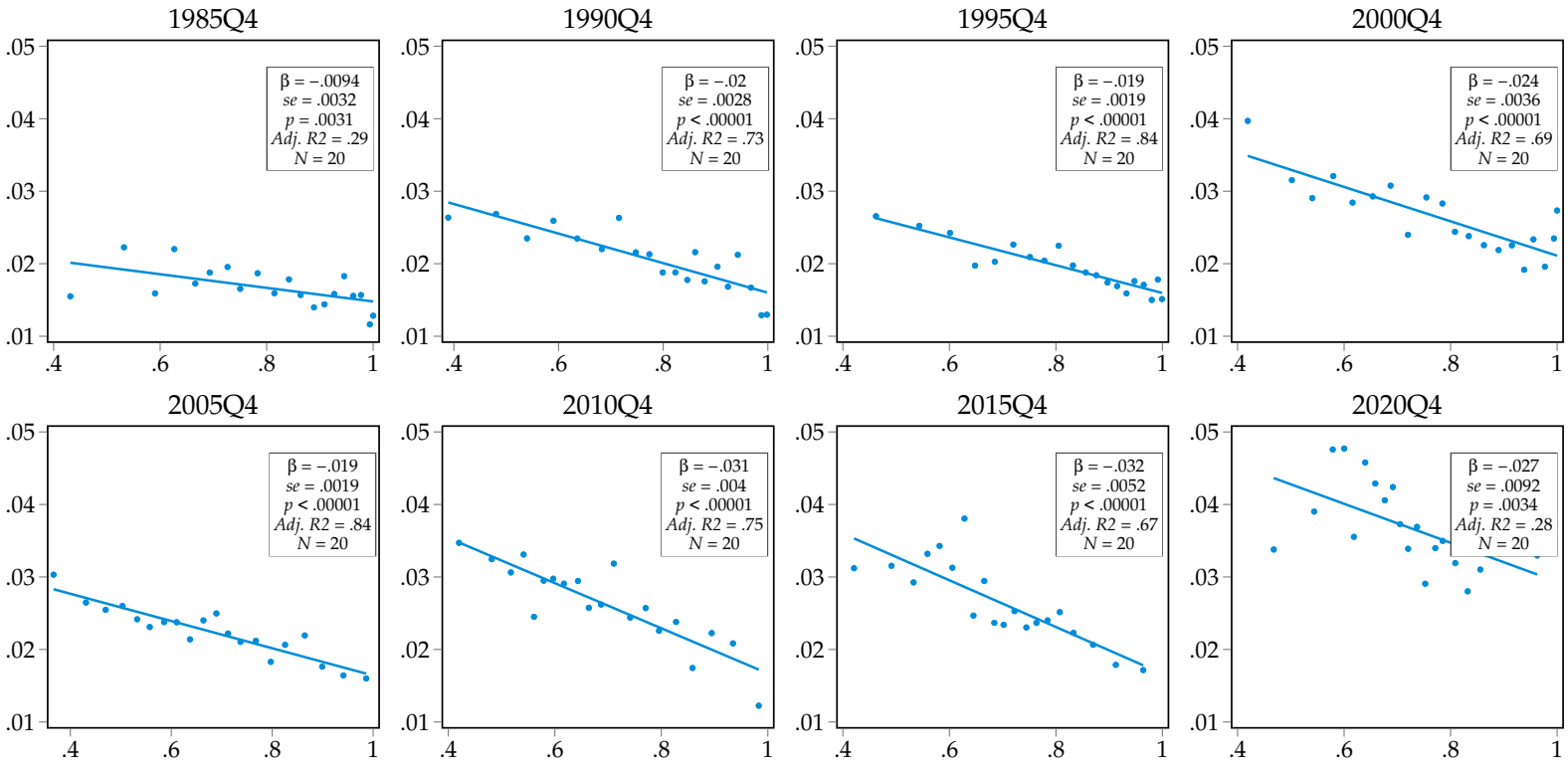
In our baseline analysis, we regress future investment rates on normalized market price (M/A), where market price is measured at the end of March and scaled by book assets. For robustness, Table C6 employs the Tobin's Q measure from Peters and Taylor (2017), which uses fiscal year-end market prices normalized by total capital. Our results remain consistent.

In unreported robustness tests, we also augment the regressions of future investment on current prices with two additional controls and their interactions with normalized market price. First, following Edmans et al. (2017), we include operating cash flow (scaled by total assets) as a non-price measure of investment opportunities. Second, per Kacperczyk et al. (2021), we control for price nonsynchronicity, calculated from the R-squared of a regression of individual stock returns on the market factor. Our results remain unchanged.



This figure shows quarter-by-quarter cross-sectional regressions of relative price informativeness (in twentiles) on trading-based ownership concentration, as measured by $ActHHI$. The estimate result reported in the first row in Panel E of Table C1 can be interpreted as a weighted average of the quarter-by-quarter slope coefficients illustrated here.

Figure C1: Relative Price Informativeness and Trading-based Ownership Concentration: HHI Index



This figure shows quarter-by-quarter cross-sectional regressions of relative price informativeness (in twentiles) on trading-based ownership concentration, as measured by $ActTop5$. The estimate result reported in the second row in Panel E of Table C1 can be interpreted as a weighted average of the quarter-by-quarter slope coefficients illustrated here.

Figure C2: Relative Price Informativeness and Trading-based Ownership Concentration: Top-5 Holdings

Table C1: Alternative Measure of Price Informativeness

This table considers several alternative measures of price informativeness. Panel A estimates price informativeness from the Post-Earnings-Announcement Drift (PEAD) model. The dependent variable is buy-and-hold abnormal returns for firm i's earnings announcement in the estimation window from day τ to day T. $Rank$ is the decile rank of the analyst earnings surprises. Panel B considers a microstructure-based measure $CPIE$ developed by Duarte et al. (2020). $CPIE$ quantifies the probability of private information arrival on a given day, derived from one of four microstructure models of private information arrival: the PIN model (PIN) of Easley et al. (1996), the adjusted PIN model (APIN) of Duarte and Young (2009), the generalized PIN model (GPIN) of Duarte et al. (2020), and the Odders-White and Ready (2008) model (OWR). Panel C considers a machine learning-based measure of informed trading intensity (ITI) by Bogousslavsky et al. (2024). ITI is trained from one of the three samples: Schedule 13D trading, opportunistic insider trades, and short sales. Panel D focuses on the q-period bias-corrected variance ratio ($VR(q)$) by Lo and MacKinlay (1988). $VR(q)$ is defined as the absolute value of the variance of returns over a q-day horizon divided by q times the variance of daily returns, minus one. We compute $VR(q)$ over horizons of q = 5, 10, 15, and 20 trading days using overlapping observations within a quarter. Panel E considers a relative price informativeness measure ($\tau_\pi^{R,j}$) by Dávila and Parlatore (2025). We divide the sample into twenty bins each quarter based on the ownership concentration of each firm ($ActHHI$ or $ActTop5$), and then aggregate the quarterly measures of relative price informativeness ($\tau_\pi^{R,j}$) within each bin-quarter. The first two rows reports the panel regression results of relative price informativeness in twentiles on the active ownership concentration variables, controlling for the quarter fixed effects. The last two rows mirror the first two rows, except that the dependent variables are the residualized form of relative price informativeness, estimated from the regression of relative price informativeness on size with quarter fixed effects. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Panel A: Post-Earnings-Announcement Drift

	$BHAR[0,2]$ (1)	$BHAR[0,2]$ (2)	$BHAR[3,24]$ (3)	$BHAR[3,24]$ (4)
$Rank$	0.0029***	0.0037***	0.0017***	0.0007
	(0.001)	(0.001)	(0.001)	(0.001)
$Rank*ActHHI$	-0.0027***		0.0015*	
	(0.001)		(0.001)	
$Rank*ActTop5$		-0.0049***		0.0016**
		(0.001)		(0.001)
$Rank*PasHHI$	0.0045***		0.0026	
	(0.001)		(0.002)	
$Rank*PasTop5$		0.0052***		0.0006
		(0.001)		(0.001)
Observations	201,240	201,240	201,240	201,240
R^2	0.172	0.173	0.150	0.150
Controls	Y	Y	Y	Y
Firm	Y	Y	Y	Y
Industry-Qtr	Y	Y	Y	Y

Continued on next page

Table C1 – *Continued*

Panel B: Conditional Probability of An Information Event								
Model:	$CPIE$ PIN (1)	$CPIE$ APIN (2)	$CPIE$ GPIN (3)	$CPIE$ OWR (4)	$CPIE$ PIN (5)	$CPIE$ APIN (6)	$CPIE$ GPIN (7)	$CPIE$ OWR (8)
$ActHHI$	-0.088***	-0.054***	-0.036***	-0.007				
	(0.012)	(0.008)	(0.012)	(0.014)				
$ActTop5$					-0.123***	-0.064***	-0.064***	-0.001
					(0.013)	(0.008)	(0.012)	(0.014)
$PasHHI$	-0.184***	-0.166***	0.055	0.168***				
	(0.034)	(0.026)	(0.044)	(0.055)				
$PasTop5$					-0.079***	-0.067***	0.020	0.090***
					(0.014)	(0.010)	(0.018)	(0.021)
Observations	66,681	66,681	66,681	66,681	66,681	66,681	66,681	66,681
R^2	0.493	0.316	0.320	0.503	0.495	0.317	0.321	0.503
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y	Y	Y
Industry-Qtr	Y	Y	Y	Y	Y	Y	Y	Y

Panel C: Informed Trading Intensity						
Training	ITI 13D (1)	ITI Insider (2)	ITI Short Sale (3)	ITI 13D (4)	ITI Insider (5)	ITI Short Sale (6)
$ActHHI$	-0.049***	-0.055***	-0.026***			
	(0.002)	(0.002)	(0.001)			
$ActTop5$				-0.057***	-0.075***	-0.035***
				(0.003)	(0.002)	(0.001)
$PasHHI$	-0.042***	-0.030***	-0.028***			
	(0.005)	(0.006)	(0.003)			
$PasTop5$				-0.028***	-0.007**	-0.017***
				(0.003)	(0.003)	(0.002)
Observations	225,723	225,653	225,754	225,723	225,653	225,754
R^2	0.329	0.291	0.471	0.333	0.303	0.482
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Qtr	Y	Y	Y	Y	Y	Y

Continued on next page

Table C1 – *Continued*

Panel D: Variance Ratio

	$VR(5)$ (1)	$VR(10)$ (2)	$VR(15)$ (3)	$VR(20)$ (4)	$VR(5)$ (5)	$VR(10)$ (6)	$VR(15)$ (7)	$VR(20)$ (8)
$ActHHI$	0.045***	0.049***	0.043***	0.036***				
	(0.003)	(0.005)	(0.006)	(0.007)				
$ActTop5$					0.045***	0.052***	0.042***	0.040***
					(0.004)	(0.006)	(0.007)	(0.008)
$PasHHI$	0.054***	0.048***	0.047***	0.042***				
	(0.006)	(0.008)	(0.010)	(0.011)				
$PasTop5$					0.058***	0.058***	0.058***	0.050***
					(0.005)	(0.006)	(0.008)	(0.009)
Observations	374,450	374,450	374,450	374,450	374,450	374,450	374,450	374,450
R^2	0.113	0.096	0.090	0.089	0.114	0.097	0.091	0.090
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y	Y	Y
Industry-Qtr	Y	Y	Y	Y	Y	Y	Y	Y

Panel E: Relative Price Informativeness

	Estimate	Std	t-value	Obs.
(1) $ActHHI$	-0.018620***	0.000412	-45.22	2960
(2) $ActTop5$	-0.020837***	0.000319	-65.22	2960
(3) $ActHHI(Residual)$	-0.003438***	0.000390	-8.82	2960
(4) $ActTop5(Residual)$	-0.001325***	0.000326	-4.07	2960

Table C2: Alternative Sample: Mutual Fund Holdings

This table replicates Tables 3-4, with the distinction that the institutional ownership data is sourced from Thomson Reuters S12 mutual fund holdings. The coefficients of the control variables are suppressed for brevity. See Table B1 for the complete list of variable definitions. Standard errors, clustered at the year and firm levels, are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Price Informativeness and $ActHHI$						
where E =	$E_{h=1}/A$ $EBIT$	$E_{h=1}/A$ $EBITDA$	$E_{h=1}/A$ NI	$E_{h=3}/A$ $EBIT$	$E_{h=3}/A$ $EBITDA$	$E_{h=3}/A$ NI
$\log(M/A)$*$ActHHI$	-0.027***	-0.034***	-0.024**	-0.049***	-0.054***	-0.028*
	(0.008)	(0.006)	(0.010)	(0.011)	(0.010)	(0.016)
$\log(M/A)$*$PasHHI$	-0.001	0.004	-0.011	0.009	0.020**	-0.008
	(0.007)	(0.005)	(0.008)	(0.010)	(0.009)	(0.015)
Observations	69,996	70,284	70,414	58,626	58,823	58,961
R^2	0.808	0.827	0.686	0.668	0.693	0.553
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y
Panel B: Price Informativeness and $ActTop5$						
where E =	$E_{h=1}/A$ $EBIT$	$E_{h=1}/A$ $EBITDA$	$E_{h=1}/A$ NI	$E_{h=3}/A$ $EBIT$	$E_{h=3}/A$ $EBITDA$	$E_{h=3}/A$ NI
$\log(M/A)$*$ActTop5$	-0.030***	-0.041***	-0.017*	-0.047***	-0.058***	-0.022
	(0.008)	(0.004)	(0.010)	(0.013)	(0.014)	(0.022)
$\log(M/A)$*$PasTop5$	0.006	0.020***	-0.011	0.011	0.041**	-0.027
	(0.009)	(0.007)	(0.011)	(0.019)	(0.016)	(0.028)
Observations	69,996	70,284	70,414	58,626	58,823	58,961
R^2	0.808	0.828	0.686	0.668	0.694	0.554
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Continued on next page

Table C2 – *Continued*

	(1)	(2)	(3)	(4)	(5)	(6)
Panel C: Investment-to-Price Sensitivity and $ActHHI$						
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
$\log(M/A)$*$ActHHI$	-0.032***	-0.022***	-0.054***	-0.049***	-0.022***	-0.072***
	(0.012)	(0.005)	(0.014)	(0.010)	(0.006)	(0.012)
$\log(M/A)$*$PasHHI$	0.024***	0.020***	0.044***	0.047***	0.028***	0.080***
	(0.006)	(0.002)	(0.007)	(0.006)	(0.006)	(0.010)
Observations	70,162	69,631	70,106	58,767	58,246	58,711
R^2	0.874	0.713	0.788	0.788	0.616	0.694
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y
Panel D: Investment-to-Price Sensitivity and $ActTop5$						
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
$\log(M/A)$*$ActTop5$	-0.031**	-0.023***	-0.053***	-0.047***	-0.017**	-0.063***
	(0.012)	(0.006)	(0.017)	(0.012)	(0.007)	(0.013)
$\log(M/A)$*$PasTop5$	0.038***	0.040***	0.078***	0.088***	0.047***	0.136***
	(0.007)	(0.006)	(0.012)	(0.010)	(0.008)	(0.017)
Observations	70,162	69,631	70,106	58,767	58,246	58,711
R^2	0.873	0.712	0.786	0.787	0.615	0.692
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Table C3: Distinguish Active/Passive Institutional Investors Using Bushee's Time-varying Classification

This table replicates Tables 3-4, with the distinction that we use Bushee's time-varying classification to distinguish active/passive institutional investors. Panels A and B focus on price informativeness, while Panels C and D focus on investment-to-price sensitivity. Trading-based ownership concentration is measured by $ActHHI$ in Panels A and C, representing the Herfindahl-Hirschman index of active institutional ownership, and by $ActTop5$ in Panels B and D, denoting the proportion of shares held by the top five active institutional investors relative to the total shares held by all active institutional investors. The sample has an annual frequency and spans from 1980 to 2022. The coefficients of the control variables are suppressed for brevity. See Table B1 for the complete list of variable definitions. Standard errors, clustered at the year and firm levels, are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Price Informativeness and $ActHHI$						
where E =	$\frac{E_{h=1}}{A}$ $EBIT$	$\frac{E_{h=1}}{A}$ $EBITDA$	$\frac{E_{h=1}}{A}$ NI	$\frac{E_{h=3}}{A}$ $EBIT$	$\frac{E_{h=3}}{A}$ $EBITDA$	$\frac{E_{h=3}}{A}$ NI
$\log(M/A)$*$ActHHI$	-0.026***	-0.030***	-0.027***	-0.052***	-0.059***	-0.035***
	(0.005)	(0.004)	(0.004)	(0.010)	(0.009)	(0.011)
Observations	83,054	83,794	83,952	69,612	70,250	70,402
R^2	0.823	0.837	0.714	0.677	0.697	0.579
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y
Panel B: Price Informativeness and $ActTop5$						
where E =	$\frac{E_{h=1}}{A}$ $EBIT$	$\frac{E_{h=1}}{A}$ $EBITDA$	$\frac{E_{h=1}}{A}$ NI	$\frac{E_{h=3}}{A}$ $EBIT$	$\frac{E_{h=3}}{A}$ $EBITDA$	$\frac{E_{h=3}}{A}$ NI
$\log(M/A)$*$ActTop5$	-0.033***	-0.040***	-0.029***	-0.045***	-0.063***	-0.022*
	(0.006)	(0.005)	(0.005)	(0.010)	(0.009)	(0.012)
Obs	83,054	83,794	83,952	69,612	70,250	70,402
R2	0.824	0.838	0.714	0.678	0.699	0.579
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Continued on next page

Table C3 – *Continued*

	(1)	(2)	(3)	(4)	(5)	(6)
Panel C: Investment-to-Price Sensitivity and $ActHHI$						
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
$\log(M/A)$*$ActHHI$	-0.022***	-0.023***	-0.044***	-0.042***	-0.025***	-0.070***
	(0.007)	(0.003)	(0.007)	(0.009)	(0.004)	(0.010)
Observations	83,616	82,913	83,549	70,174	69,462	70,098
R^2	0.863	0.692	0.771	0.765	0.613	0.681
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y
Panel D: Investment-to-Price Sensitivity and $ActTop5$						
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
$\log(M/A)$*$ActTop5$	-0.027***	-0.024***	-0.050***	-0.049***	-0.028***	-0.082***
	(0.009)	(0.005)	(0.012)	(0.011)	(0.004)	(0.013)
Observations	83,616	82,913	83,549	70,174	69,462	70,098
R^2	0.863	0.693	0.771	0.766	0.616	0.684
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Table C4: Concentration Measures Based on Trading Volume

This table replicates Tables 3-4, with the distinction that trading volume, rather than holdings, is utilized to construct measures of trading-based ownership concentration. Specifically, concentration is measured by $ActHHI$ in Panels A and C, representing the Herfindahl-Hirschman index of trading volume of active institutional investors, and by $ActTop5$ in Panels B and D, denoting the proportion of trading volume of the top five active institutional investors relative to the total trading volume of all active institutional investors. Panels A and B focus on price informativeness, while Panels C and D focus on investment-to-price sensitivity. The sample has an annual frequency and spans from 1980 to 2022. The coefficients of the control variables are suppressed for brevity. See Table B1 for the complete list of variable definitions. Standard errors, clustered at the year and firm levels, are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Price Informativeness and $ActHHI$						
where E =	$E_{h=1}/A$ $EBIT$	$E_{h=1}/A$ $EBITDA$	$E_{h=1}/A$ NI	$E_{h=3}/A$ $EBIT$	$E_{h=3}/A$ $EBITDA$	$E_{h=3}/A$ NI
$\log(M/A)$*$ActHHI$	-0.013***	-0.016***	-0.015**	-0.033***	-0.035***	-0.029***
	(0.005)	(0.004)	(0.005)	(0.010)	(0.009)	(0.010)
Observations	84,042	84,783	84,944	70,442	71,084	71,236
R^2	0.821	0.835	0.712	0.673	0.693	0.576
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y
Panel B: Price Informativeness and $ActTop5$						
where E =	$E_{h=1}/A$ $EBIT$	$E_{h=1}/A$ $EBITDA$	$E_{h=1}/A$ NI	$E_{h=3}/A$ $EBIT$	$E_{h=3}/A$ $EBITDA$	$E_{h=3}/A$ NI
$\log(M/A)$*$ActTop5$	-0.024***	-0.027***	-0.027***	-0.035***	-0.041***	-0.028***
	(0.005)	(0.004)	(0.007)	(0.008)	(0.008)	(0.009)
Observations	84,042	84,783	84,944	70,442	71,084	71,236
R^2	0.821	0.835	0.712	0.673	0.694	0.576
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Continued on next page

Table C4 – *Continued*

	(1)	(2)	(3)	(4)	(5)	(6)
Panel C: Investment-to-Price Sensitivity and $ActHHI$						
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
$\log(M/A)$*$ActHHI$	-0.017***	-0.014***	-0.031***	-0.025***	-0.011***	-0.037***
	(0.006)	(0.003)	(0.009)	(0.007)	(0.004)	(0.008)
Observations	84,607	83,891	84,538	71,004	70,282	70,928
R^2	0.862	0.689	0.768	0.761	0.607	0.675
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y
Panel D: Investment-to-Price Sensitivity and $ActTop5$						
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
$\log(M/A)$*$ActTop5$	-0.012*	-0.008*	-0.021**	-0.014*	-0.003	-0.021*
	(0.007)	(0.004)	(0.010)	(0.008)	(0.005)	(0.010)
Observations	84,607	83,891	84,538	71,004	70,282	70,928
R^2	0.862	0.688	0.768	0.762	0.609	0.676
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Table C5: Concentration Measures Without Distinguishing Active/Passive Investors

This table replicates Tables 3-4, with the distinction that we reconstruct the trading-based concentration measures without distinguishing active/passive investors. Ownership concentration among all institutional investors is measured by $TotHHI$ in Panels A and C, and $TotTop5$ in Panels B and D. Specially, $TotHHI_{i,q} = \frac{\sum_{j=1}^{N_{tot}} (S_{i,j,q}^2)}{(\sum_{j=1}^{N_{tot}} S_{i,j,q})^2}$ captures HHI of institutional shares, where N_{tot} denotes the number of institutions holding stock i; $TotTop5_{i,q} = \frac{\sum_{j=1}^{\text{Top 5}} S_{i,j,q}}{\sum_{j=1}^{N_{tot}} S_{i,j,q}}$ measures the proportion of shares held by the top five largest institutional investors relative to the total shares held by all institutional investors. Panels A and B focus on price informativeness, while Panels C and D focus on investment-to-price sensitivity. The sample has an annual frequency and spans from 1980 to 2022. The coefficients of the control variables are suppressed for brevity. See Table B1 for the complete list of variable definitions. Standard errors, clustered at the year and firm levels, are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Price Informativeness and $TotHHI$						
where E =	$\frac{E_{h=1}}{A}$ $EBIT$	$\frac{E_{h=1}}{A}$ $EBITDA$	$\frac{E_{h=1}}{A}$ NI	$\frac{E_{h=3}}{A}$ $EBIT$	$\frac{E_{h=3}}{A}$ $EBITDA$	$\frac{E_{h=3}}{A}$ NI
$\log(M/A)$*$TotHHI$	-0.045***	-0.046***	-0.047***	-0.081***	-0.078***	-0.075***
	(0.006)	(0.006)	(0.009)	(0.016)	(0.014)	(0.014)
Observations	108,681	109,667	109,876	91,018	91,869	92,074
R^2	0.805	0.816	0.694	0.659	0.675	0.569
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y
Panel B: Price Informativeness and $TotTop5$						
where E =	$\frac{E_{h=1}}{A}$ $EBIT$	$\frac{E_{h=1}}{A}$ $EBITDA$	$\frac{E_{h=1}}{A}$ NI	$\frac{E_{h=3}}{A}$ $EBIT$	$\frac{E_{h=3}}{A}$ $EBITDA$	$\frac{E_{h=3}}{A}$ NI
$\log(M/A)$*$TotTop5$	-0.039***	-0.041***	-0.037***	-0.066***	-0.066***	-0.052***
	(0.005)	(0.004)	(0.006)	(0.010)	(0.010)	(0.009)
Observations	108,681	109,667	109,876	91,018	91,869	92,074
R^2	0.805	0.816	0.694	0.660	0.676	0.569
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Continued on next page

Table C5 – *Continued*

	(1)	(2)	(3)	(4)	(5)	(6)
Panel C: Investment-to-Price Sensitivity and $TotHHI$						
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
$\log(M/A)$*$TotHHI$	-0.020***	-0.021***	-0.040***	-0.033***	-0.033***	-0.068***
	(0.006)	(0.005)	(0.011)	(0.009)	(0.007)	(0.012)
Observations	109,450	108,359	109,304	91,771	90,687	91,613
R^2	0.847	0.643	0.730	0.738	0.570	0.641
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y
Panel D: Investment-to-Price Sensitivity and $TotTop5$						
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
$\log(M/A)$*$TotTop5$	-0.015***	-0.018***	-0.032***	-0.018**	-0.021***	-0.041***
	(0.006)	(0.004)	(0.009)	(0.007)	(0.006)	(0.011)
Observations	109,450	108,359	109,304	91,771	90,687	91,613
R^2	0.847	0.643	0.730	0.738	0.571	0.643
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Table C6: Alternative Measure of Investment-to-Price Sensitivity

This table replicates Table 4, but replaces the normalized market price ($\frac{M}{A}$) with the Tobin's Q measure from Peters and Taylor (2017). Q is calculated as fiscal year-end market prices normalized by total capital (taking the logarithm). Trading-based ownership concentration is measured by $ActHHI$ in Panel A, representing the Herfindahl-Hirschman index of active institutional ownership, and by $ActTop5$ in Panel B, denoting the proportion of shares held by the top five active institutional investors relative to the total shares held by all active institutional investors. The coefficients of the control variables are suppressed for brevity. See Table B1 for the complete list of variable definitions. Standard errors, clustered at the year and firm levels, are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: $ActHHI$			
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
Q*ActHHI	-0.025***	-0.029***	-0.051***	-0.051***	-0.027***	-0.076***
	(0.007)	(0.003)	(0.008)	(0.009)	(0.005)	(0.011)
Observations	79,099	78,456	79,040	66,505	65,845	66,432
R^2	0.873	0.696	0.782	0.770	0.615	0.686
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y
			Panel B: $ActTop5$			
where I=	$I_{h=1}/K$ $Intangible$	$I_{h=1}/K$ $Physical$	$I_{h=1}/K$ $Invest$	$I_{h=3}/K$ $Intangible$	$I_{h=3}/K$ $Physical$	$I_{h=3}/K$ $Invest$
Q*ActTop5	-0.035***	-0.036***	-0.069***	-0.066***	-0.035***	-0.102***
	(0.010)	(0.005)	(0.013)	(0.011)	(0.006)	(0.013)
Observations	79,099	78,456	79,040	66,505	65,845	66,432
R^2	0.874	0.696	0.783	0.772	0.619	0.690
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

D Alternative Explanations

In this section, we consider and rule out several alternative explanations for the negative association between active institutional ownership concentration and price informativeness. For brevity, we report only the results for the 1-year prediction horizon.

D.1 Short-sale Constraints

One alternative interpretation is that stocks with more concentrated ownership coincide with higher short-sale constraints. In this case, arbitrageurs may refrain from correcting mispricing, leading to low informational efficiency. For instance, Porras Prado et al. (2016) claim that investors with larger holdings are reluctant to lend stocks due to concerns that short selling could decrease stock prices and weaken their monitoring control. The authors find that stocks with more concentrated ownership exhibit lower lending supply and higher shorting costs.

To evaluate this alternative explanation, we combine equity lending data sourced from Markit with our main sample. The combined sample shrinks in size because equity lending data is only available from 2002 onward. If the negative effect is mainly driven by short-sale constraints rather than investors' learning and trading decisions, it should diminish when stocks are easy to borrow. Following Muravyev et al. (2022), a stock is considered easy to borrow if the indicative borrowing fee is less than or equal to 1%. We exclude stocks with annual average indicative borrowing fee greater than 1% and re-conduct our baseline analysis in the remaining easy-to-borrow subsample. Simultaneously, we include lending supply ($Supply$) along with its interaction term with market price in the regression models. Lending supply is defined as the annual average dollar value of lendable shares relative to a firm's market capitalization.

Panels A and B of Table D1 report the results. The negative effect of active ownership concentration on price informativeness remains significant in the easy-to-borrow subsample with the additional control for lending supply. Therefore, it is unlikely that the documented results are

attributed to the short-sale constraints.

D.2 Institutional Price Pressure

Another alternative interpretation is that stocks with more concentrated ownership face higher institutional price pressure (Greenwood and Thesmar, 2011). For example, forced fire sales by mutual funds experiencing large outflows can create large nonfundamental price declines in underlying stocks. Stocks with more concentrated ownership are likely more vulnerable to such nonfundamental price pressure, resulting in lower informational efficiency.

To examine this possibility, we construct a measure of mutual fund fire-sale pressure, $Pressure$, following Wardlaw (2020). Formally,

$$Pressure_{i,t} = \sum_{j}^{m} \frac{|F_{j,t}|}{TA_{j,t-1}} \times \frac{Shares_{i,j,t-1}}{Shrout_{i,t-1}},$$

where $F_{j,t}$ denotes the net dollar flow to mutual fund j in quarter t, $TA_{j,t-1}$ is the total asset value of mutual fund j in quarter $t-1$, $Shares_{i,j,t-1}$ indicates the shares of firm i held by fund j in quarter $t-1$, and $Shrout_{i,t-1}$ denotes the total shares outstanding of firm i in quarter $t-1$. In the calculation, only funds with outflow greater than 5% of total assets ($\frac{F_{j,t}}{TA_{i,j,t-1}} < -5\%$) are included because they are most likely to be forced into a fire sale of their holdings. This measure captures the total share amount of each stock sold by funds with extreme outflows, if all of the funds were to sell their stocks in proportion to their initial holdings. We aggregate $Pressure$ to the firm-year level by averaging its quarterly values, and assign a value of zero to firms with missing values.

Panels C and D of Table D1 replicate the baseline regressions, but explicitly control for $Pressure$ and its interaction term with the market price. The effect of institutional price pressure on informational efficiency is not evident. While the effect is negative when examining real investment, as the coefficients of $\log(M/A) * Pressure$ are negative and weakly significant in Panel D, it is overturned in the price informativeness case in Panel C. In contrast, the coefficients on the inter-

action term between our concentration measures and market price remain significantly negative across all specifications. Therefore, our main results do not seem to be driven by institutional price pressure.

D.3 Voluntary Disclosure

Another possibility is that the negative impact on informational efficiency stems from companies' voluntary disclosure decisions. When active ownership concentration increases, large shareholders become more influential. They may gain easier access to managers and substitute private communication for public information acquisition. Consequently, managers have less incentive to provide voluntary disclosures, leading to higher information asymmetry and lower informational efficiency. To explore this possibility, we investigate whether firms with higher active ownership concentration issue fewer management forecasts, a proxy for voluntary disclosure.

We obtain management forecasts from the LSEG IBES guidance database, available since 1993. Following prior studies (e.g., Boone and White, 2015), we treat all management forecasts issued on the same day as a single observation, as they are usually released in one press release. We also exclude pre-earnings announcement forecasts made on or after the fiscal period end, as per previous research (e.g., Rogers and Stocken, 2005). Our first measure of voluntary disclosure is guidance frequency ($GuideFreq$), representing the number of management forecasts at the firm-year level. This includes all types of quarterly and annual forecasts, such as earnings, sales, and capital expenditure. Firms without forecasts in a given year receive a value of zero. The second measure is $GuideDummy$, which indicates the propensity of voluntary disclosure and equals one if any management forecasts are provided. We also focus on management earnings forecasts, which are the most common and notable guidance. $EPSGuideFreq$ and $EPSGuideDummy$ measure the frequency and propensity of earnings forecasts.

Panels E and F of Table D1 report regressions of year $t+1$ management forecasts on year t ownership concentration. Beyond baseline controls and fixed effects, we include book-to-market

ratio, ROA, and log book assets as per prior studies (e.g., Chen et al., 2008). If concentrated ownership hinders voluntary disclosure, we would expect negative coefficients for our concentration measures ($ActHHI$ and $ActTop5$). This is not the case. Across all specifications, the coefficients are either insignificant or even significantly positive. Thus, reduced voluntary disclosure does not seem to explain our results.

Table D1: Alternative Explanations

This table considers and rules out several alternative explanations for the negative association between active institutional ownership concentration and informational efficiency. Panels A and B replicate Tables 3-4 in the easy-to-borrow subsample, with additional controls for lending supply along with its interaction term with market price. A stock is considered easy to borrow if its annual average indicative borrowing fee is less than or equal to 1%. Lending supply, $Supply$, is defined as the annual average dollar value of lendable shares relative to a firm's market capitalization. Panels C and D replicate Tables 3-4 while controlling for institutional price pressure along with its interaction term with market price. Panels E and F regress voluntary disclosure in year $t+1$ on active ownership concentration in year t. We construct four proxies of voluntary disclosure: $GuideFreq$ denotes the number of management forecasts at the firm-year level. This includes all types of quarterly and annual forecasts, such as earnings, sales, and capital expenditure. Firms without forecasts in a given year receive a value of zero. $GuideDummy$ indicates the propensity of voluntary disclosure and equals one if any management forecasts are provided. $EPSGuideFreq$ and $EPSGuideDummy$ measure the frequency and propensity of earnings forecasts. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Panel A: Short Sale Constraints and Price Informativeness

where E =	$E_{h=1}/A$ $EBIT$ (1)	$E_{h=1}/A$ $EBITDA$ (2)	$E_{h=1}/A$ NI (3)	$E_{h=1}/A$ $EBIT$ (4)	$E_{h=1}/A$ $EBITDA$ (5)	$E_{h=1}/A$ NI (6)
$\log(M/A)$*$ActHHI$	-0.017**	-0.019***	-0.017**			
	(0.006)	(0.006)	(0.006)			
$\log(M/A)$*$ActTop5$				-0.031***	-0.033***	-0.027***
				(0.005)	(0.005)	(0.007)
$\log(M/A)$*$Supply$	-0.001	-0.004	0.028***	0.003	-0.000	0.032***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Observations	34,962	34,954	34,965	34,962	34,954	34,965
R^2	0.821	0.829	0.699	0.821	0.829	0.699
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Panel B: Short Sale Constraints and Investment-to-Price Sensitivity

where I=	$I_{h=1}/K$ $Intangible$ (1)	$I_{h=1}/K$ $Physical$ (2)	$I_{h=1}/K$ $Invest$ (3)	$I_{h=1}/K$ $Intangible$ (4)	$I_{h=1}/K$ $Physical$ (5)	$I_{h=1}/K$ $Invest$ (6)
$\log(M/A)$*$ActHHI$	-0.006**	-0.007**	-0.016***			
	(0.003)	(0.003)	(0.005)			
$\log(M/A)$*$ActTop5$				-0.010***	-0.010***	-0.022***
				(0.003)	(0.003)	(0.005)
$\log(M/A)$*$Supply$	-0.018***	-0.004	-0.024***	-0.018***	-0.004	-0.023***
	(0.005)	(0.005)	(0.008)	(0.005)	(0.004)	(0.008)
Observations	34,841	34,799	34,838	34,841	34,799	34,838
R^2	0.932	0.769	0.840	0.932	0.769	0.840
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Continued on next page

Table D1 – *Continued*

Panel C: Liquidity Shock and Price Informativeness						
where E =	$E_{h=1}/A$ $EBIT$ (1)	$E_{h=1}/A$ $EBITDA$ (2)	$E_{h=1}/A$ NI (3)	$E_{h=1}/A$ $EBIT$ (4)	$E_{h=1}/A$ $EBITDA$ (5)	$E_{h=1}/A$ NI (6)
$\log(M/A)$*$ActHHI$	-0.026***	-0.030***	-0.027***			
	(0.005)	(0.004)	(0.004)			
$\log(M/A)$*$ActTop5$				-0.033***	-0.040***	-0.028***
				(0.006)	(0.005)	(0.005)
$\log(M/A)$*$Pressure$	0.017*	0.011	0.054***	0.019	0.012	0.056***
	(0.010)	(0.009)	(0.012)	(0.012)	(0.011)	(0.014)
Observations	83,054	83,794	83,952	83,054	83,794	83,952
R^2	0.823	0.837	0.714	0.824	0.838	0.714
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y
Panel D: Liquidity Shock and Investment-to-Price Sensitivity						
where I=	$I_{h=1}/K$ $Intangible$ (1)	$I_{h=1}/K$ $Physical$ (2)	$I_{h=1}/K$ $Invest$ (3)	$I_{h=1}/K$ $Intangible$ (4)	$I_{h=1}/K$ $Physical$ (5)	$I_{h=1}/K$ $Invest$ (6)
$\log(M/A)$*$ActHHI$	-0.022***	-0.023***	-0.044***			
	(0.007)	(0.003)	(0.007)			
$\log(M/A)$*$ActTop5$				-0.027***	-0.024***	-0.050***
				(0.009)	(0.005)	(0.012)
$\log(M/A)$*$Pressure$	-0.013	-0.015*	-0.024	-0.015	-0.016*	-0.027*
	(0.012)	(0.009)	(0.015)	(0.012)	(0.008)	(0.014)
Observations	83,616	82,913	83,549	83,616	82,913	83,549
R^2	0.863	0.692	0.771	0.863	0.693	0.771
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y

Continued on next page

Table D1 – *Continued*

Panel E: Concentration and Voluntary Disclosure (Management Forecast)				
	$GuideFreq$		$GuideDummy$	
	(1)	(2)	(3)	(4)
$ActHHI$	-0.259		0.005	
	(0.350)		(0.025)	
$ActTop5$		0.016		0.050*
		(0.427)		(0.026)
Observations	72,223	72,223	72,223	72,223
R^2	0.660	0.660	0.628	0.628
Controls	Y	Y	Y	Y
Firm	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y
Panel F: Concentration and Voluntary Disclosure (Management Earnings Forecast)				
	$EPSGuideFreq$		$EPSGuideDummy$	
	(1)	(2)	(3)	(4)
$ActHHI$	-0.107		0.011	
	(0.144)		(0.022)	
$ActTop5$		0.087		0.072***
		(0.180)		(0.024)
Observations	72,223	72,223	72,223	72,223
R^2	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Firm	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y

E Competing Hypothesis: Corporate Governance

Our theoretical model and empirical findings indicate that active ownership concentration impairs informational efficiency. However, a competing hypothesis posits that concentration enhances shareholder engagement and corporate governance, potentially improving informational efficiency. For instance, Hartzell and Starks (2003) document a positive relationship between institutional ownership concentration and the pay-for-performance sensitivity of executive compensation. Furthermore, prior research (Cai et al., 2006; Lee et al., 2016) demonstrates that informational efficiency increases with governance quality. Thus, concentrated ownership could theoretically enhance informational efficiency through better governance. The negative effects identified in our baseline analysis suggest that this potential governance channel, if present, is dominated by the learning and information pass-through mechanisms. In this section, we empirically assess the validity of this competing hypothesis by examining whether concentrated ownership is assoc iated with improved corporate governance.

Following prior research (e.g., Albuquerque et al., 2022; Guo et al., 2021), we measure shareholder activism using Schedule 13D filings.[6] We obtain Schedule 13D filings from the Audit Analytics Shareholder Activism database, which has been available since 2000. Audit Analytics categorizes each filing into one of seven types: concern, dispute, control, discussion, support, agreement, and other. Agrawal and Lim (2022) classify filings in the categories of concern, dispute, and control as hard activism, characterized by confrontational and hostile tactics. Based on this classification, we create two dummy variables: $Activism$, which is set to one if a firm has a 13D filing in any category during a given year; and $HardActivism$, which is set to one if a firm has a 13D filing in the categories of concern, dispute, or control within a given year.

In addition, we use corporate fraud as a proxy for the quality of corporate governance. Fewer cases of corporate fraud indicate better corporate governance. Following Liu (2016), we create a fraud dummy variable ($Fraud$), which is equal to one if the firm experiences any of the fol-

[6] Gaining 5% or more ownership in a public company triggers an SEC filing obligation. The choice between filing the detailed Schedule 13D or the simpler Schedule 13G depends on the investor's intentions: those aiming to actively influence company management file a 13D, while passive investors opt for a 13G.

lowing three events in a given year. First, the firm is subject to class action lawsuits in a given year. Securities class action lawsuit filings are sourced from the Stanford Securities Class Action Clearinghouse (SCAC). Following previous studies (e.g., Kempf and Spalt, 2023), cases related to initial public offering underwriter allocation, analyst coverage, and mutual funds, rather than firm management, are excluded. Moreover, to ensure the materiality of the cases, we exclude those that are subsequently dismissed, as per Dyck et al. (2010). Second, earnings are misstated in that firm-year according to the SEC's Accounting and Auditing Enforcement Releases (AAER), which are issued for violations of SEC Rule 10b-5. AAER data is sourced from Dechow et al. (2011). Third, the firm announced an earnings restatement in that year according to the Audit Analytics Database. We exclude restatements arising from clerical errors (identified by the Audit Analytics "RES_CLER_ERR" flag) to distinguish material financial restatements from unintentional errors (Wang, 2022). We further require the restatements disclosed via Form 8-K Item 4.02, as these are deemed more material due to SEC regulatory requirements (Cahan et al., 2024).

Table E1 reports regression results of the aforementioned proxies on active ownership concentration. We retain the same control variables and fixed effects as in the baseline model. The coefficients on the concentration measures ($ActTop5$ and $ActHHI$) are statistically insignificant and sensitive to different specifications. When we use $HardActivism$ to measure shareholder activism, the coefficient of $ActTop5$ is significantly positive. However, this result is not robust: the coefficient of the alternative concentration measure, $ActHHI$, is insignificant and economically close to zero. Therefore, we find no empirical support for the competing hypothesis that active ownership concentration enhances informational efficiency through improved corporate governance.

Table E1: Competing Hypothesis: Corporate Governance

This table reports regression results of shareholder activism and corporate fraud proxies on trading-based ownership concentration. $Activism$ is set to one if a firm has a 13D filing during a given year; $HardActivism$ is set to one if a firm has a 13D filing in the categories of concern, dispute, or control within a given year. $Fraud$ is set to one if the firm experienced any of the following events in a given year: involvement in class action lawsuits, earnings misstated according to the SEC's AAER, or an earnings restatement per the Audit Analytics. Trading-based ownership concentration is measured by $ActHHI$, representing the Herfindahl-Hirschman index of active institutional ownership, and by $ActTop5$, denoting the proportion of shares held by the top five active institutional investors relative to the total shares held by all active institutional investors. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

	$Activism$		$HardActivism$		$Fraud$	
	(1)	(2)	(3)	(4)	(5)	(6)
$ActHHI$	-0.003		0.004		-0.011	
	(0.020)		(0.009)		(0.007)	
$ActTop5$		0.015		0.021**		-0.000
		(0.019)		(0.009)		(0.009)
Observations	52,368	52,368	52,368	52,368	88,256	88,256
R^2	0.408	0.409	0.210	0.210	0.353	0.353
Controls	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y