When ChatGPT Stops Talking:

GenAI-Induced Retail Herding and Systematic Risk

Xiaoxue He

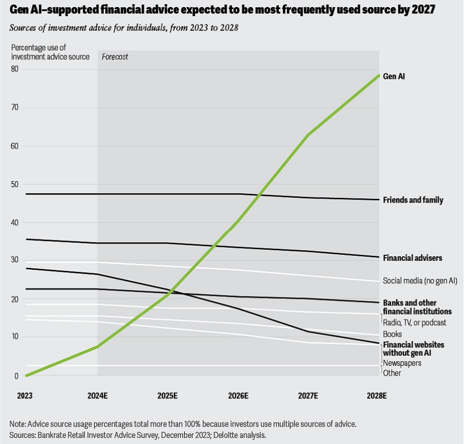
Lancaster University Management School

4 June 2026

11th BoF/ESRB Conference on AI and Systemic Risk Analytics

The GenAI Adoption Wave

- $>$**40%** of retail investors use GenAI for investment decisions *(eToro 2023; Oliver Wyman 2024)*
- GenAI will become the **dominant source of retail investment advice** by 2027, reaching **78%** of users by 2028 *(Deloitte 2024)*.



Motivation

Unclear Net Effect of GenAI adoption on Retail Investors

- **Well-documented Bright Side:** Emerging evidence suggests that GenAI can foster more informative retail trading (e.g. Cheng et al., 2025; Ecker et al., 2026)
- **Largely Unexplored Dark Side:** Regulators warn that GenAI may amplify investors' behavioral biases and create a new systematic risk channel, especially "GenAI-induced herding" *(ESMA, 2023; BIS, 2024; FSB, 2024; IMF, 2024; Bank of England, 2025)*, yet empirical evidence remains scarce.

GenAI-induced herding

Applying centralized GenAI architecture
⇒ compressing diverse information sets into homogenized outputs
⇒ synchronizing individual beliefs and increasing correlated trades
⇒ transforming idiosyncratic retail noise into a coordinated market force.

Research Questions and Main Findings

Main Question

Does GenAI-induced herding exist, and what are its consequences for market stability?

1. **[GenAI-induced herding] Does ChatGPT availability increase retail synchronicity?**
   - ChatGPT serves as a "belief homogenizer" within and across stocks, producing similar trading decisions.
2. **[Good or Bad Herding?] Does synchronized trading reflect improved price discovery?**
   - It reflects aggregated overreaction, with short-run predictability followed by reversal.
3. **[Market Cost] What are the market consequences?**
   - Liquidity worsens and systematic volatility increases.

Theoretical Foundation

1. **Retail investors as "noise traders"**
   - Retail investors have long been characterized as "noise traders" with ingrained behavioral biases (Barber and Odean, 2008; Barber et al., 2009).
2. **Overreaction hypothesis**
   - Retail investors have a tendency to overreact (e.g., De Bondt and Thaler, 1985; Daniel et al., 1998; Barber and Odean, 2008).
   - Perceived information advantage is a key driver of overreaction (Liu et al., 2022).
3. **Lee, Shleifer & Thaler (1991) noise-trader model**
   - If individual noise is independent, its risk would be diversifiable;
   - if noise becomes correlated, the resulting "noise trader risk" becomes systematic.

Contributions

1. **New evidence on the net effect of GenAI on financial markets**

   Prior work: GenAI improves individual-level participation and informativeness (e.g. Cheng et al., 2025; Ecker et al., 2026).

   - This study sheds light on the dark side of GenAI by investigating the effect of GenAI adoption on investors' collective behavioral biases and its market-level consequences.

2. **Contribute to the behavioral finance literature**

   - **Novel form of herding: "GenAI-induced herding"**

     Distinct from traditional social learning or imitation-based herding, "GenAI-induced herding" arises from the centralized architecture of the technology itself.

   - **Modern empirical application of Lee, Shleifer & Thaler (1991) noise trader model**

     This study documents that GenAI converts diversifiable idiosyncratic noise to systematic risk.

3. **Early empirical evidence for regulatory concerns**

   GenAI-induced herding is increasingly flagged as a top capital-markets risk of broader GenAI adoption, yet empirical evidence remains scarce *(e.g. IMF, 2024; Bank of England, 2025)*.

   - This study provides early empirical evidence that GenAI-induced herding introduces new systematic risk.

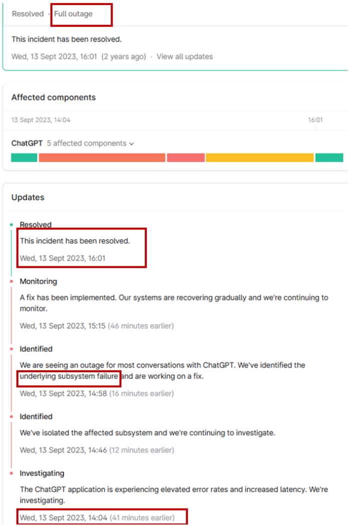
Identification Strategy: Unexpected ChatGPT Outages

Why ChatGPT?

- No tech prerequisite; publicly accessible
- Accounting for 77% of all GenAI traffic (Liu et al., 2025).
- **Best approximation of the average GenAI user** (Chatterji et al., 2025).

21 Full ChatGPT Outages during 2023–2024

- Sourced from OpenAI's official incident reports (https://status.openai.com/)
- Incidents that are classified as full outages
- Overlap with U.S. equity trading hours (9:30–16:00 ET)



Identification Strategy: Unexpected ChatGPT Outages

Why ChatGPT?

- No tech prerequisite; publicly accessible
- Accounting for 77% of all GenAI traffic (Liu et al., 2025).
- **Best approximation of the average GenAI user** (Chatterji et al., 2025).

21 Full ChatGPT Outages during 2023–2024

- Sourced from OpenAI's official incident reports (https://status.openai.com/)
- Incidents that are classified as full outages
- Overlap with U.S. equity trading hours (9:30–16:00 ET)

Matched-Window Design

Treatment window: all 5-min intervals within the outage periods

Control: same clock-time 5-min intervals from prior 5 trading days, same stock

$\Rightarrow$ Absorbs: time-invariant firm characteristics + intraday seasonality

$\Rightarrow$ Isolates: causal effect of ChatGPT access

Data and Sample

Data Sources

- **Millisecond intraday trading data**: NYSE's TAQ database
  - → **Retail trade identification:** BJZZ algorithm (Boehmer et al. (2021); Barber et al. (2024))
- **Stock prices & daily returns**: CRSP
- **Firm characteristics**: Compustat
- **Firm-specific information events**: Capital IQ Key Developments
- **Analyst coverage**: I/B/E/S

Sample Construction

	Obs.	Stocks
Initial sample*	5,682,191	4,201
Less: Low retail activity	(3,448,427)	(1,155)
Less: <10 retail trades	(525,460)	(354)
Less: Missing controls	(146,611)	(352)
Final sample	**1,561,693**	**2,340**

Initial Sample:

- Intraday 5-minute intervals for U.S. ordinary common stocks listed on NYSE, NYSE American, or Nasdaq (excluding non-ordinary shares and sub-$1 stocks) during documented outage and matched control periods, from 2023 to 2024.

Finding 1: ChatGPT Raises Retail Participation

Baseline Specification:

$$\text{Retail Intensity}_{i,t,p} = \alpha + \beta_1 \text{Outage}_{t,p} + \gamma \text{Controls}_{i,t} + \varepsilon$$

- *Retail Intensity*$_{i,t,p}$: retail trading volume (count) during 5-minute interval *p* for stock *i* on day *t*
- *Outage*$_{t,p}$: indicator that equals one if interval *p* on day *t* falls within a ChatGPT outage window, and zero otherwise
- *Controls*$_{i,t}$: control for past returns (overnight, prior-day, and days t-5 to t-2), stock price, market capitalization, book-to-market ratio, intraday volatility, and analyst coverage
- *FEs*: firm + event×interval
- *SE*: clustered by firm and event×date

	Volume (Mean: 0.79)	**Trade Count** (Mean: 65.2)
Outage	-0.044^{**} (-2.30)	-4.100^{**} (-2.40)
% Change	**−5.6%**	**−6.3%**

Takeaway

Retail investors are meaningfully using ChatGPT in trading process.

Robustness:

- Excludes firm news days ✓
- Excludes CPI/PPI/FOMC days ✓
- Excludes outages <30 min ✓
- Excludes pre-10AM outages ✓
- Placebo test: insignificant

Finding 2: ChatGPT as a "Belief Homogenizer"

Within-Stock Opinion Divergence

	MATO	ΔTO	SUV
Outage	**0.107*** **	**0.103*** **	**0.064*** **
	(3.10)	(3.14)	(2.66)

Cross-Stock Coordination

	Outage	Controls	Diff
PC1_Ind	0.141	0.208	**-0.067*** **
PC1_Market	0.114	0.170	**-0.056*** **

Takeaway

- ChatGPT facilitates belief homogenization within and across stocks.

Net Retail Order Flow

	Abs_ROI	ROIV
Outage	0.467	**0.006*** **
	(1.36)	(2.92)

Takeaway

- ChatGPT-driven belief homogenization can translate into similar trading decisions.

Finding 3a: Synchronized Trading Reflects Aggregated Overreaction

Return Predictability of Retail Order Imbalance

- **Price discovery**: informed trading produce persistent return predictability.
- **Overreaction**: overreaction-driven trading should produce short-lived predictability followed by reversal

	+5min	+10min	+15min	+25min	+1 day	+3 day	+5 day
Outage x Selling ROI	**-0.013***	**-0.022***	**-0.030***	**-0.041***	**0.544***	**0.484***	**0.290**
	(-1.81)	(-1.83)	(-1.87)	(-1.68)	(1.88)	(1.87)	(1.08)
Outage x Buying ROI	0.007	0.002	-0.002	0.001	-0.156	-0.545***	-0.284
	(1.21)	(0.18)	(-0.13)	(0.02)	(-0.73)	(-2.79)	(-1.19)
Selling ROI	**0.009*****	**0.012*****	**0.010****	**0.007**	**-0.092**	**-0.196****	**-0.191****
	(3.67)	(2.90)	(2.08)	(0.94)	(-1.53)	(-2.44)	(-1.99)
Buying ROI	0.007***	0.012***	0.019***	0.019*	0.081	0.160	0.239*
	(2.68)	(2.81)	(2.97)	(1.79)	(1.14)	(1.50)	(1.92)

Takeaway

Synchronized retail selling facilitated by ChatGPT reflects aggregated overreaction rather than informative price discovery.

Finding 3b: Ruling Out Other Channels for Return Reversals

Return Predictability of Decomposed Retail Order Imbalance

- **ROI Decomposition**: Follow Boehmer et al. (2021) and decompose selling ROI into three components: **ROI persistence** (a proxy for price pressure channel), **ROI liquidity** (a proxy for liquidity provision channel), and **ROI other** (residual component including behavioral channel).

	+5min	+10min	+15min	+25min	+1 day	+3 day	+5 day
Outage x ROI_other	**-0.011**	**-0.023***	**-0.026***	**-0.040***	**0.517***	**0.670*****	**0.486***
	(-1.49)	(-1.67)	(-1.68)	(-1.84)	(1.92)	(2.79)	(1.93)
Outage x ROI_liquidity	-0.004	-0.001	-0.012	-0.049	-0.009	-0.141	0.122
	(-0.32)	(-0.06)	(-0.31)	(-0.79)	(-0.05)	(-0.50)	(0.42)
Outage x ROI_persistence	-0.003	-0.013	-0.022	-0.034	1.057**	0.733	0.349
	(-0.26)	(-0.71)	(-0.86)	(-0.84)	(2.53)	(1.36)	(0.65)
ROI_other	**0.010*****	**0.012*****	**0.012****	**0.014***	**-0.095**	**-0.160****	**-0.142**
	(3.56)	(2.72)	(2.11)	(1.75)	(-1.58)	(-2.10)	(-1.66)
ROI_liquidity	0.001	0.000	-0.006	-0.003	0.191	0.483**	0.564**
	(0.11)	(0.02)	(-0.22)	(-0.09)	(1.51)	(2.10)	(2.52)
ROI_persistence	0.005	0.007	0.001	-0.004	-0.001	-0.151	0.116
	(0.95)	(0.77)	(0.08)	(-0.19)	(-0.01)	(-0.91)	(0.56)

Finding 3c: Information Load Attenuates Overreaction

Intuition:
With more information available, ChatGPT is less likely to generate homogeneous outputs, so the overreaction effect should weaken.

Triple Interaction: Outage $\times$ SellROI $\times$ Information Load

Information Load $=$ log(1 $+$ # firm-specific events in Capital IQ, prior 30 days)

	+5min	+10min	+15min	+25min	+1 day	+3 day	+5 day
Panel A. Selling ROI							
Outage x ROI x Information_Load	0.012***	0.020**	0.023**	0.047**	-0.627***	-0.565**	-0.476*
	(3.27)	(2.45)	(2.07)	(2.43)	(-3.62)	(-2.38)	(-1.66)
Panel B. Decomposed Selling ROI							
Outage x ROI_other x Information_Load	0.012***	0.015*	0.01	0.027**	-0.478***	-0.290*	-0.331
	(2.77)	(1.97)	(1.12)	(2.18)	(-4.41)	(-1.77)	(-1.59)

Triple-interaction signs are opposite to the baseline Outage $\times$ ROI effect
$\Rightarrow$ Outage effect attenuated for Information-rich firms.

Finding 4a: ChatGPT Harms Market Liquidity

Illiquidity Falls During Outages:

	Price Impact	Quoted Spread	Effective Spread
Outage	-0.003^{**}	-0.003^{**}	-0.004^{**}
	(-2.03)	(-2.50)	(-2.17)
% Change	**−15%**	**−2.5%**	**−4.4%**

A Possible Mechanism: Crowding Out Contrarians

	ROI		
	Full Sample	Buying	Selling
Outage x Return_Overnight	**-0.014*** **	**-0.005*** **	**-0.009*** **
	(-3.04)	(-3.12)	(-2.86)
Return_Overnight	-0.003***	-0.001***	-0.002***
	(-4.81)	(-4.91)	(-3.97)

Takeaway

ChatGPT availability may dampen retail investors' contrarian trading by directing attention to common signals, thereby reducing liquidity provision.

Finding 4b: ChatGPT Elevates Systematic Volatility

Volatility Falls During Outages

Outage	Coeff.	% Change
Panel A: 5-min trade-based volatility		
volatility	-0.298^{**}	**−9.0%**
Panel B: Idiosyncratic & Systematic Volatility		
Firm-specific	-0.087^{***}	**−12.1%**
Industry-level	-0.037^{***}	**−20.6%**
Market-level	-0.066^{***}	**−27.5%**

Lee, Shleifer & Thaler (1991) noise-trader model

If individual noise becomes correlated across traders and across assets, the resulting "noise trader risk" is non-diversifiable and can be priced.

Takeaway

ChatGPT affects volatility primarily by synchronizing idiosyncratic noise, thereby increasing systematic risk.

Conclusion

What We Do

Exploit 21 unexpected ChatGPT outages (2023–2024) as exogenous shocks to study how GenAI adoption affects retail investors' collective behavioral biases and market quality.

Three Core Findings:

1. **ChatGPT is a belief homogenizer**: the availability of ChatGPT reduces within-stock opinion divergence; increases cross-stock price synchronicity; stabilizes net retail order flow
2. **Aggregated overreaction instead of improved price discovery**: short-run return predictability followed by reversal on the sell side when ChatGPT is available.
3. **Market-wide costs are real and systematic**: the availability of ChatGPT worsens liquidity; rises systematic risk by synchronizing idiosyncratic noise.

Limitations

1. **Single platform & potential substitution concern:** Evidence is from ChatGPT only. If investors switch to other GenAI tools during outages, estimates likely understate the broader GenAI effect.
2. **No prompt/response data:** Queries or outputs can not be directly observed; cannot pin down informational channels behind belief homogenization.
3. **No account-level data:** No account-level trades data; can't directly connect an individual's GenAI use to synchronized trading/overreaction.
4. **Retail-active stocks only:** Focus on stocks with meaningful retail activity; effects may differ in less retail-dominated segments.

Thank you!

Email: x.he21@lancaster.ac.uk

When ChatGPT Stops Talking: GenAI-induced Retail Herding and Systematic Risk

Xiaoxue He

Department of Accounting and Finance, Lancaster University Management School, Lancaster, LA1 4YX, UK

x.he21@lancaster.ac.uk

THIS VERSION: MARCH 2026

ABSTRACT

Using 21 unexpected ChatGPT outages as exogenous shocks to GenAI availability, we study how GenAI adoption affects retail investors' collective behavioral biases and market quality. We find that ChatGPT acts as a "belief homogenizer," synchronizing retail beliefs and trading decisions. This synchronization reflects collective overreaction rather than a collective-level improvement in information processing. ChatGPT availability worsens market liquidity by crowding out retail investors' natural contrarian liquidity-provision role. We further show that ChatGPT threatens market stability by aggregating idiosyncratic retail noise into undiversifiable systematic risk.

Keywords: Generative AI; ChatGPT; Retail Investors; Behavioral Bias

1 Introduction

Generative artificial intelligence (GenAI) has rapidly democratized complex financial information processing, significantly lowering the barriers to informed market participation. Publicly available GenAI platforms such as ChatGPT have great potential to reshape how retail investors interact with financial markets, particularly for those traditionally constrained by high information costs and limited access to sophisticated analytical tools. Survey evidence already points to meaningful adoption among retail investors: over 40% of retail investors report using GenAI tools to inform their investment decisions (eToro, 2023; Blankespoor et al., 2024; Oliver Wyman, 2024), and projections suggest these applications will become the dominant source of retail investment advice by 2027, reaching 78% of users by 2028 (Deloitte Center, 2024). Retail investors have long been characterized as "noise traders" whose market participation is hampered by significant informational disadvantages (Lee, 1992; Blankespoor et al., 2020) and deeply ingrained behavioral biases (Barber and Odean, 2008; Barber et al., 2008). The speed of GenAI's diffusion into this noise population raises a fundamental question that has not yet been fully answered: what does widespread AI adoption actually do to retail investors and the market they participate in?

Emerging research has begun to document the individual-level effects of GenAI on investors, with broadly positive findings. Several studies find that GenAI tools foster retail trading activity and support more informed investment decisions (e.g., Chang et al., 2023; Cheng et al., 2025; Even-Tov et al., 2025; Ecker et al., 2026; Hansen and Lee, 2025). However, this optimistic individual-level narrative overlooks a potentially consequential dark side: GenAI's capacity to amplify retail investors' behavioral biases. At the individual level, GenAI may reinforce individual-level overconfidence as reflected in their illusions of knowledge, precision, and control, leading to excessive and lower quality retail trades (e.g., Barber and Odean, 2002; Croom, 2025; Havakhor et.al., 2025). At the collective level, the concern is potentially more consequential. Because GenAI models operate through a centralized architecture, they risk compressing diverse information sets into homogenized outputs and delivering similar signals to all users simultaneously. This creates a structural mechanism for synchronizing individual beliefs and increasing correlated retail order flow, thereby aggregating idiosyncratic retail noise into a material market force. If these concerns are founded, the documented increase in GenAI-associated retail trading may reflect synchronized collective overreaction rather than genuine improvements in market efficiency.

Regulators have already flagged this as a burgeoning systemic risk, referring to it as "GenAI-induced herding" (ESMA, 2023; BIS, 2024; FSB, 2024; IMF, 2024; Bank of England, 2025). This motivates our investigation into whether GenAI-induced herding exists and what its consequences are for market stability.

The behavioral foundation for this concern is well established. A large literature documents retail investors' tendency to overreact (e.g., De Bondt and Thaler, 1985; Daniel et al., 1998; Barber and Odean, 2008), and recent evidence suggests that a key driver of overreaction is perceived information advantage (Liu et al., 2022), which may be strengthened by information technologies such as GenAI tools (Havakhor et al., 2025). The aggregate market impact of individual-level overreaction is often limited as long as investors make independent mistakes that partially offset each other. However, the consequences become systemic when retail errors are correlated across investors or across assets simultaneously, because the resulting risk is no longer diversifiable (Shleifer and Summers, 1990; Lee et al., 1991; Drake et al., 2017). This is at the heart of our study. We argue that GenAI introduces a structurally distinct mechanism for generating correlated noise rapidly at scale: individual beliefs and trading behavior can be synchronized through shared exposure to a centralized information-processing technology and its homogenized signals, rather than via the social learning and imitation in traditional herding models (Bikhchandani et al., 1992; Devenow and Welch, 1996). To our knowledge, this mechanism and its implications for market stability have not been examined empirically.

To study these questions, we exploit 21 unexpected ChatGPT full outages between 2023 and 2024 that overlap U.S. equity trading hours as plausibly exogenous shocks to GenAI availability for retail investors. ChatGPT is the natural focus of our study. It offers a publicly accessible and user-friendly interface without tech knowledge pre-requirement, making it the GenAI tool of choice for the broad retail investor population. It surpassed 100 million users within two months of its launch in November 2022 - a pace of adoption widely regarded as unprecedented for consumer technology (Reuters, 2023). By mid-2025, 700 million users were sending 18 billion messages per week, representing approximately 10% of the global adult population (Reuters, 2025; Forbes, 2025). ChatGPT has also maintained its market dominance despite intensifying competition, accounting for 77% of traffic across the top 60 GenAI tools as of April 2025 (Liu et al., 2025). Because ChatGPT's user base dwarfs that of any competing platform, we expect our sample to be a closer approximation to the behavior of the average GenAI user than any alternative tool could offer (Chatterji et al., 2025).

We combine these outage events with high-frequency trade and quote data from the NYSE TAQ database and identify retail trades following Boehmer et al. (2021) and Barber et al. (2024). For each outage, we compare outcomes within the outage window to the same clock-time 5-minute intervals from the five preceding trading days for the same stock. This matched-window design absorbs time-invariant firm characteristics and intraday seasonality, allowing us to isolate how retail behavior and market quality change when access to ChatGPT is disrupted.

Our analysis centers on three main questions: Does ChatGPT availability increase the synchronicity of retail trading, consistent with GenAI-induced herding? Does this synchronized trading reflect collective-level improvement in information processing or aggregated overreaction? And does such synchronization translate into market-wide costs in the form of weaker liquidity or higher volatility?

We begin by establishing that ChatGPT availability meaningfully raises retail trading intensity. Retail trading volume falls by 5.6% and retail trade count by 6.3% during outages, confirming that ChatGPT meaningfully engages in retail trading decision-making and providing the empirical foundation for the analyses that follow.

We then find evidence that ChatGPT shapes the collective beliefs of retail investors. At the stock level, within-stock opinion divergence, measured by market-adjusted turnover, abnormal turnover, and scaled unexplained volume following Garfinkel (2009), increases significantly during outages, consistent with ChatGPT acting as a belief homogenizer across investors. At the industry and market levels, price synchronicity falls significantly during outages, and principal component analysis shows that the industry- and market-level first principal components of 5-minute returns explain 6.7 and 5.6 percentage points less variance during outages than during control periods, respectively. These results indicate that ChatGPT facilitates belief homogenization not only across investors within individual stocks but also across stocks simultaneously. We further demonstrate that this belief homogenization translates into more correlated retail trading decisions, as reflected in a steadier net retail order flow (lower volatility of retail order imbalance) during control periods.

Third, we examine the key question for market efficiency: whether ChatGPT-driven synchronization reflects collective-level information processing improvement or aggregated overreaction. We distinguish between informed trading and behavioral overreaction by examining the temporal profile of return predictability. Prior literature suggests that informed trading should exhibit persistent return predictability, reflecting a permanent price impact. Conversely, if excessive trading driven by overreaction would show a pattern of short-term continuation followed by a reversal as prices revert to fundamental values (De Bondt and Thaler, 1985; Hong and Stein, 1999). We examine the predictive relationship between retail order imbalance (*ROI*) and future returns at intraday (5–25 minutes) and multi-day (1–5 days) horizons. On the sell side, we find a clear pattern of return reversal: selling ROI predicts returns effectively in the short term but reverses at 3–5 day horizons, consistent with investor overreaction rather than permanent price discovery. Meanwhile, the interaction term between selling ROI and the outage indicator is generally opposite in sign to the baseline coefficient, indicating that the reversal is attenuated when ChatGPT is unavailable. This pattern is concentrated on the sell side because selling decisions are constrained by existing positions and therefore more tightly linked to deliberate information processing, making them a cleaner test of GenAI's effect (Barber and Odean, 2008; Barber et al., 2022). We further follow Boehmer et al. (2021) and decompose selling ROI into components reflecting price pressure, liquidity provision, and residual selling pressure. The reversal and its outage-related attenuation are concentrated entirely in the residual component, ruling out compensation for liquidity provision as the source of the observed reversal (Barber et.al, 2008; Barrot et.al, 2016; Barber et.al, 2022) and strengthening the behavioral overreaction interpretation. We also show that the outage effect on predictability is weaker for firms with higher information load, measured by the number of firm-specific events in Capital IQ Key Developments, consistent with the intuition that ChatGPT is less likely to generate homogeneous trading suggestions across users with abundant available information.

Finally, we document the market consequences of this aggregated overreaction. It has been well established in behavioral asset pricing literature that the aggregate forces generated by investors' systematic behavioral biases can harm market stability, represented by higher inventory risk faced by liquidity providers and elevated volatility (e.g. Odean, 1998; Eaton et.al.,2022; Bogousslavsky and Collin-Dufresne, 2023). Consistent with this mechanism, we find that liquidity improves significantly during outages: price impact falls by 15% and effective spreads narrow by 4.4%. We trace this improvement to retail investors reverting to their natural contrarian strategies that supply liquidity to the market when ChatGPT is unavailable. Our results also show that intraday volatility falls by around 9% during outages. Decomposing volatility into firm-specific, industry-level, and market-level components following Campbell et al. (2001), we find that the largest proportional changes occur in systematic components: industry- and market-level volatility drop by 20.6% and 27.5%, respectively, compared to 12.1% for firm-specific volatility. This decomposition indicates that ChatGPT's dominant effect on market stability operates through coordination channel in which ChatGPT increases co-movement across stocks. This pattern is consistent with the Lee et al. (1991) mechanism: ChatGPT converts what would be diversifiable idiosyncratic retail noise into correlated, systematic price pressure that cannot be diversified away.

Our paper makes several contributions. First, we contribute to the growing literature on GenAI and financial markets. Existing work has focused primarily on whether GenAI tools improve retail investors' individual-level market participation and information processing (Cheng et al., 2025; Ecker et al., 2026; Even-Tov et al., 2025; Chang et al., 2023; Hansen and Lee, 2025). We shed light on the dark side of GenAI by investigating the effect of AI adoption on investors' collective behavioral biases and its market-level consequences, which is central to understanding the net welfare effects of AI in financial markets and has not been previously documented.

Second, we contribute to the behavioral finance literature by identifying a novel mechanism of belief synchronization. Unlike traditional herding, which often relies on social learning or mimicking, "GenAI-induced herding" arises from the centralized architecture of the technology itself. We show that when millions of investors utilize the same underlying model, idiosyncratic interpretations are compressed into a homogenized signal. This finding provides a modern empirical application of the Lee et.al (1991) noise-trader model: we document how GenAI aggregates diversifiable idiosyncratic noise into a systematic risk factor.

Finally, our findings provide early empirical evidence for the regulatory concern that widespread GenAI adoption may create herding behavior and elevate systematic risk in financial markets (ESMA, 2023; BIS, 2024; FSB, 2024; IMF, 2024; Bank of England, 2025). We document measurable belief synchronization, aggregated overreaction, and market quality deterioration arising from a single dominant GenAI platform. Because our estimates are derived from outages of that single platform, they likely represent a conservative lower bound: to the extent that investors partially substitute to competing tools during outages, the true effect of AI availability is larger than we document. This also sheds light on the risks of AI market concentration and adds a new empirical dimension to the urgent regulatory agenda around AI in finance.

The remainder of the paper proceeds as follows. Section 2 discusses the relevant literature. Section 3 describes the data and sample construction. Section 4 presents empirical results. Section 5 concludes.

2 Literature Review

2.1 Retail Behavioral Biases: Herding and Overreaction

A large behavioral finance literature establishes that individual investors are unsophisticated noise traders prone to behavioral biases that push prices away from fundamentals (e.g., Barber and Odean, 2000; Kumar and Lee, 2006; Barber and Odean, 2008; Barber et al., 2008). Two biases are central to our study: herding and overreaction.

2.1.1 Herding

Theoretical models of herding divide broadly into rational and irrational classes. Rational herding models emphasize reputational concerns and information aggregation. Reputation-based models show that agents may rationally mimic peers to protect their reputational standing or hedge evaluation risk (Scharfstein and Stein, 1990; Trueman, 1994; Graham, 1999). Information-cascade models show how herding can arise when investors rationally infer information from others' actions and downweight their own private signals (Banerjee, 1992; Bikhchandani et al., 1992). Froot et.al (1992) show that short-horizon speculators may herd on common information to anticipate what other informed investors know, even when that information is unrelated to fundamentals. Shiller (2006) further point out that mass dissemination of information may create similar thinking among large groups of people.

Irrational herding models attribute correlated investor behavior to sentiment and biased beliefs rather than deliberate imitation. Shleifer and Summers (1990) argue that irrational noise traders respond to pseudo-signals such as broker recommendations or popular narratives, rather than fundamentals, generating systematic demand distortions. De Long et.al (1990) further shows that noise traders with erroneous and correlated beliefs create persistent mispricing that rational arbitrageurs cannot fully correct.

Empirical herding research has largely focused on institutions, but a growing set of studies examines retail herding. Barber et al. (2008) document systematic patterns in individual investor herding that are difficult to reconcile with standard institutional mechanisms such as principal-agent incentives. More recent work using retail order imbalance as a direct measure of correlated retail trading confirms that herding exists among individual investors (e.g., Barber et al., 2022; Eaton et al., 2022; Barber et al., 2024b).

2.1.2 Overreaction

The overreaction literature begins with De Bondt and Thaler (1985), who document long-run return reversals consistent with investors overweighting recent performance and extrapolating trends beyond what fundamentals justify. Daniel et al. (1998) provide a formal model in which overconfidence and biased self-attribution generate price pressure that eventually reverses as prices revert. These frameworks imply a clear empirical signature: informed trading should produce persistent return predictability, while overreaction-driven trading should produce short-lived predictability followed by reversal (e.g., Barber & Odean, 2008; Kaniel et al., 2008; Kremer & Nautz, 2013; Hillert et al., 2014). Liu et al. (2022) provide direct survey evidence from retail investors confirming that perceived information advantage is a key motivation for excessive retail trading.

2.2 Noise Trader Risk: From Idiosyncratic Noise to Priced Systematic Risk

Whether retail behavioral biases are merely individually costly or are also collectively harmful to market stability depends on a single critical question: are investor errors independent or correlated across traders and assets?

De Long et. al (1990) develop the seminal framework for this question, showing that irrational noise traders with stochastic sentiment can persistently move prices away from fundamental values. However, if noise traders' misperceptions to individual assets are uncorrelated, arbitrageurs would eliminate any possible mispricing. Shleifer and Summers (1990) synthesize and extend this noise trader approach, emphasizing that arbitrage is both costly and risky, and that the capacity of noise trader demand to distort prices depends critically on whether that demand is correlated across investors: when noise trading is dispersed, it is easier for the market to absorb; when noise trading is aligned, it creates aggregate shifts in demand that are harder to offset. Lee et.al (1991) further enhance the importance of correlation structure of noise: if noise trading were independent across assets and investors, its risk would be diversifiable; if fluctuations are correlated across traders and across many assets, the resulting “noise trader risk” is non-diversifiable and can be priced.

2.3 Generative AI and Financial Markets: Emerging Evidence

The literature on generative AI in accounting and finance is rapidly expanding. Existing research focuses on the natural language processing capabilities of ChatGPT, for example, generating historical business summaries from annual reports (Breitung and Müller, 2025), facilitating analyst information production (Bertomeu et.al., 2025; Bradshaw et.al., 2025), and predicting stock market reactions from news headlines (Lopez-Lira and Tang, 2023).

A second and more directly relevant strand examines how GenAI affects investor trading behavior. Blankespoor et al. (2024) provide survey evidence on retail investors' use and perceptions of GenAI for investment decisions. Cheng et al. (2025), the paper closest to our setting, use unexpected ChatGPT outage design to show that ChatGPT increases retail trading volume and facilitates informed trading. Ecker et al. (2026) analyze large-scale GenAI query data from a major provider and document that GenAI usage is associated with more informed trading, lower liquidity, and that aggregated answer sentiment correlates with same-day abnormal returns. Chang et al. (2023) observe a significant increase in the AI-sentiment alignment of retail traders following the wide deployment of ChatGPT. Even-Tov et al. (2025) exploit a temporary ban on ChatGPT in Italy as an exogenous shock to GenAI access and find that the ban increases correlation across holdings and raises portfolio volatility without affecting abnormal performance.

There are also several laboratory experiments that shed light on GenAI's effect on investors behavioral biases. For example, Croom (2025) argues that GenAI interactivity blurs the attribution of skill and increases perceived ability, strengthening overconfidence-like behavior. Hansen and Lee (2025) use large language model agents to replicate classic herding experiments, finding that although AI agents rely more heavily on private information than humans, they can still be induced to herd when explicitly guided toward profit-maximizing decisions. However, to our knowledge, empirical evidence is still limited.

3 Data and Sample

We identify ChatGPT outage events in 2023–2024 from OpenAI's publicly disclosed incident reports, which provide the official records of date, severity, and incident timeline for each service disruption.[1] We focus on incidents classified as full outages, during which

[1] See https://status.openai.com/history for OpenAI's incident history.

ChatGPT is broadly inaccessible to users. For each outage, we collect the start and end timestamps in Eastern Time (ET) and retain events whose incident window overlaps at least partially with U.S. equity trading hours (9:30–16:00 ET). Appendix A provides a complete list of outage events used in our analysis. The final event set contains 21 full outages that occur on different days and at different times, with durations ranging from 16 to 390 minutes (2,120 minutes in total; mean duration of about 101 minutes).

Our intraday trade and quote data come from the NYSE's TAQ database. We obtain daily stock prices and returns from CRSP, firm characteristics from Compustat, firm-specific information events from Capital IQ Key Developments, and analyst coverage from I/B/E/S.

We begin with U.S. ordinary common stocks listed on NYSE, AMEX, or Nasdaq and exclude non-ordinary shares and stocks priced below $1. Because our setting is retail-focused, we identify retail trades (purchases and sales) using the approach in Boehmer et al. (2021) and Barber et al. (2024a), which exploits institutional features of retail execution.[2] Retail orders are often executed off-exchange and receive small price improvement relative to the national best bid and offer (NBBO). Specifically, we classify retail trades as those reported with TAQ exchange code "D" (off-exchange) and with prices 0.1–0.4 cents below a round penny for buys and 0.1–0.4 cents above a round penny for sells. We do not classify trades priced exactly at a round penny or near the half-penny (0.4–0.6 cents, inclusive) to reduce misclassification.

Following prior work (e.g., Eaton et al., 2022; Cheng et al., 2025), we aggregate retail transactions to 5-minute intervals. For each outage, we include all 5-minute observations that fall within the outage window. We construct a matched control sample using the same clock time 5-minute intervals from the previous five trading days for the same stock. We further exclude any control-day window that overlaps with another ChatGPT outage.[3]

We then apply additional screens. Referring to Eaton et al. (2022), we exclude stocks with limited retail activity by dropping those with fewer than 500 average daily retail trades during the control period. Since our main tests rely on retail order imbalance, we also drop 5-minute

[2] BJZZ's classification approach has been widely adopted in recent studies of retail trading (e.g., Eaton et al., 2022; Cheng et al., 2025; Havakhor et al., 2025).

[3] For example, the ChatGPT outage on March 20, 2023 ran from 12:41 to 22:54 ET. Since the outage extends beyond market close, we truncate the window at 16:00 ET, yielding 40 5-minute intervals within outage window on March 20. The matched control sample covers the same 40 intervals on each of the five preceding trading days (March 13, 14, 15, 16, and 17, 2023). We exclude March 15 from the control window because a separate ChatGPT outage occurred on that day.

intervals with fewer than 10 retail trades (Barber et al., 2024b). Finally, we exclude observations with missing control variables.

The final sample contains 1,561,693 stock–date–5-minute observations from 2,340 unique stocks. Table 1 summarizes the sample selection process.

[Insert Table 1 here]

Table 2 reports summary statistics for the main sample. To minimize the influence of outliers, we winsorize all continuous variables at the 1st and 99th percentiles, except stock returns. Variable definitions are provided in Appendix B. The table shows that retail trading is economically active at the 5-minute level: a typical interval contains around 65 retail trades, corresponding to roughly 7,900 shares per stock-interval on average. We also noticed that market quality measures (spreads, price impact, and volatility) are heavily right-skewed, consistent with the wide dispersion observed in retail order imbalance and investor disagreement proxies.

[Insert Table 2 here]

4 Research Design and Empirical Results

4.1 The Effect of ChatGPT Outages on Retail Trading Intensity

We begin by examining whether retail trading activity declines when investors temporarily lose access to ChatGPT. If ChatGPT meaningfully supports retail trading decisions, we expect a decrease in retail trading activity during outage periods. We estimate:

$$Retail\ trading\ intensity_{i,t,p} = \alpha + \beta_1 Outage_{t,p} + \gamma Controls_{i,t} + \varepsilon_{i,\mathrm{t},\mathrm{p}} \qquad (1)$$

where i, t, and p indicate stock, date, and 5-min interval, respectively. We measure retail trading intensity using *Retail Trading Volume* and *Retail Trade Count*. The independent variable, *Outage t,p*, equals one if interval *p* on day *t* falls within a ChatGPT outage window. and zero otherwise. The coefficient *β1* captures the difference in retail trading intensity between outage intervals and the matched control intervals for a stock. Following prior work (e.g., Eaton et al., 2022; Cheng et al., 2025), we control for past returns (overnight, prior-day, and days t-5 to t-2), stock price, market capitalization, book-to-market ratio, intraday volatility, and analyst coverage. We include firm fixed effects and event-by-interval fixed effects to absorb time-invariant firm heterogeneity and time-varying market-wide conditions within each event and time-of-day bin. Standard errors are two-way clustered at the firm and event-by-date levels.

Table 3 shows that retail trading activity declines significantly when ChatGPT is unavailable, consistent with Cheng et al. (2025). Specifically, retail trading volume falls by 0.044 relative to a sample mean of 0.79, implying a 5.6% reduction (t= -2.30), and the retail trading count experiences a 6.3% decrease (=4.100/ 65.20; t= -2.40). These estimates suggest that access to ChatGPT meaningfully raises retail participation even at short intraday horizons.

[Insert Table 3 here]

A primary identification concern is whether the documented outage effects are truly exogenous or driven by confounding market factors. We address this in two steps. First, we identify the root causes of outages from OpenAI's incident reports and find no systematic evidence that outages are related to market conditions.[4] Second, we conduct a battery of robustness tests to ensure our results are not spurious, summarized in Table 4.

[Insert Table 4 here]

To rule out the influence of information shocks, Panel A excludes observations coinciding with firm-specific news events. The results are essentially unchanged (Retail Trading Volume: −0.044, t = −2.30; Retail Trade Count: −4.238, t = −2.35). Panel B drops events during which scheduled macroeconomic announcements, specifically US Consumer Price Index (CPI), Producer Price Index (PPI), and the Federal Open Market Committee (FOMC) minutes releases, are made before or during the outage window, as these announcements are known to trigger significant market reactions.[5] While four events are removed from the sample, the results remain similar (−0.043, t = −2.02; −3.912, t = −2.06).[6] Panel C excludes outages shorter than 30 minutes to rule out the effect of minor disruptions, and the results

[4] While OpenAI does not disclose root causes of all outages, we are still able to identify the causes of 11 events based on the incident narratives. These include database instability, node provisioning failures, underlying subsystem failures, software bugs, authentication issues, cloud provider data-center power failures, and deployment changes. None of these causes are plausibly related to stock market conditions.

[5] Prior research documents that CPI, PPI, and FOMC announcements are major macroeconomic news which can lead to strong stock market reaction (e.g., Savor and Wilson, 2013; Lucca and Moench, 2015; Fisher et al., 2022). CPI and PPI announcement dates are obtained from the Bureau of Labor Statistics (https://www.bls.gov/bls/archived_sched.htm), and FOMC minutes dates are from the Federal Reserve Board (https://www.federalreserve.gov/monetarypolicy/fomccalendars.htm).

[6] The four excluded events are: February 14, 2023 (CPI announcement), March 15, 2023 (PPI announcement), May 24, 2023 (FOMC minutes release), and September 13, 2023 (CPI announcement).

hold (-0.045, t = -2.25; -4.190, t = -2.34).[7] To mitigate potential confounding from after-hours news or elevated opening-period volatility, Panel D drops outages beginning before 10:00 AM, and the results remain statistically and economically significant (−0.043, t = −2.23; −6.066, t = −2.53).[8] Finally, Panel E presents a placebo test by assigning the outage indicator to day t−1 rather than the actual outage day. The placebo coefficients are statistically insignificant (t = −1.34; t = −1.64), confirming that the observed outage effect is unique to actual outage periods and not a mechanical artifact of the research design. We report similar robustness checks for subsequent analyses in Appendix C. Taken together, these tests substantially mitigate identification concerns.

One additional concern is that retail investors may partially switch to competing AI tools, such as Google Gemini or Anthropic's Claude, during ChatGPT outages.[9] To the extent substitution occurs, our estimates represent a conservative lower bound on the true effect, implying the real effect may be even more significant than our empirical results suggest. More broadly, this speaks to the risks of concentration in AI tools: if a single dominant platform shapes the beliefs and trading behavior of millions of retail investors, any technical disruption or systematic bias in its outputs could have outsized effects on financial markets. Our findings thus also provide early empirical support for regulatory attention to AI concentration risk (e.g., ESMA, 2023; BIS, 2024; FSB, 2024; IMF, 2024; Bank of England, 2025).

4.2 The Effect of ChatGPT Outages on Retail Investors' Collective Beliefs

Having established that ChatGPT raises retail participation, we now investigate whether it also shapes the collective beliefs of retail investors. If ChatGPT acts as a centralized information-processing tool, it should reduce belief heterogeneity by delivering a common informational signal to users, thereby directing retail investors toward similar trading decisions. We examine this conjecture along two dimensions: within-stock belief homogenization and cross-stock coordination.

4.2.1 Within-Stock Belief Homogenization

[7] Three outage events are excluded: March 15 and May 9, 2023, and October 17, 2024.
[8] Five outage events are excluded: February 21, February 27, August 31, September 13, and November 8, 2023.
[9] We attempted to conduct a parallel analysis using outage events for these competing platforms but found that neither maintains a complete and reliable public incident history comparable to OpenAI's structured status page.

We start by testing whether ChatGPT affects retail opinion divergence at the stock level. Following Garfinkel (2009), we use market-adjusted turnover *(MATO)*, abnormal market-adjusted turnover *(ΔTO),* and standardized unexplained volume *(SUV)* as proxies for within-stock differences in trading views. [10] We estimate:

$$Opinion\ Divergence_{i,t,p} = \alpha + \beta_1 Outage_{t,p} + \gamma Controls_{i,t} + \varepsilon_{i,t,p} \tag{2}$$

Table 5 shows that all three measures of opinion divergence increase significantly during outage periods, with *MATO* rising by 0.107 (t = 3.10), *ΔTO* by 0.103 (t = 3.14), and *SUV* by 0.064 (t = 2.66). These results indicate that ChatGPT unavailability raises within-stock opinion divergence among retail investors. Put differently, the pattern is consistent with ChatGPT acting as a "belief homogenizer" that dampens heterogeneous interpretations across investors of individual stocks.

[Insert Table 5 here]

4.2.2 Cross-Stock Coordination

We extend this analysis to cross-stock coordination. Because ChatGPT synthesizes both firm-specific and broader industry and market information when generating investment insights, it may also foster coordination across stocks. We explore this using two approaches.

First, we construct price synchronicity (*SYNCH*) following Piotroski and Roulstone (2004) at three levels: within-industry, market-wide, and overall systematic, as a measure of cross-stock return co-movement. We estimate:

$$SYNCH_{i,t,p} = \alpha + \beta_1 Outage_{t,p} + \gamma Controls_{i,t} + \varepsilon_{i,t,p} \tag{3}$$

Panel A of Table 6 shows that price synchronicity decreases significantly at all three levels during outage periods (within-industry: -0.219, t= -1.76; market: -0.251, t=-1.96; −0.179, t= −1.87), suggesting that ChatGPT availability promotes stock co-movements broadly. [11]

[Insert Table 6 here]

[10] These proxies for investor opinion divergence have been widely used in the literature (e.g., Lou et al., 2014; Holzman et al., 2021; Watkins, 2022).

[11] Our main tests use the Fama–French 12-industry classification and we require the industry-interval contains at least 10 firm. The findings still hold when using the Fama–French 30- and 48-industry classifications.

Second, following Kaniel et al. (2008), we conduct industry-level and market-level principal component analysis (PCA) of 5-minute interval returns and *ROI* separately for outage and matched control windows, and compare the amount of variance explained by the leading principal components. We describe the PCA procedure in detail in Appendix D. The leading principal components capture the strongest common movements across stocks; a larger share of variance explained by these components indicates greater cross-stock coordination (Hasbrouck and Seppi, 2001). Therefore, if ChatGPT drives cross-stock coordination, the leading components are expected to explain more variance during control periods than during outages. Panel B of Table 6 confirms this. For example, within industries, the first principal component explains 20.8% of the variation in 5-minute returns during control periods, compared to only 14.1% during outage periods, a statistically significant gap of 6.7%; at the market level, the first principal component explains 5.6% more variance during control periods than during outages. We find a consistent pattern in the PCA of *ROI* at the industry level, while the market-level *ROI* result is directionally consistent but not statistically significant.[12]

Taken together, the evidence in Sections 4.2.1 and 4.2.2 establishes that ChatGPT homogenizes retail beliefs both within and across stocks simultaneously, creating the precondition for diversifiable idiosyncratic retail noise to become correlated, systematic, and priced.

4.2.3 One-Sidedness and Stability of Net Retail Order Flow

We next examine whether ChatGPT-driven belief homogenization translates into similar trading decisions, as reflected in the one-sidedness and stability of net retail order flow. If ChatGPT coordinates retail trades, it may push order flow to be more one-sided on average, and more likely, it should make net order flow more stable over time by reducing back-and-forth swings between net buying and net selling that arise from idiosyncratic disagreement. We therefore expect retail order flow to become less one-sided but more volatile during outages.

We measure one-sidedness using retail order imbalance (*ROI*). Following Bradley et al. (2024), trade-based *ROI* assigns relatively more weight to smaller trades and thus better

[12] Our PCA results are comparable to Kaniel et al. (2008), who find that the first principal component of returns explains 12.07% of the variance and the first principal component of individual investor order imbalances explains 1.70%.

captures the behavior of smaller retail investors, while volume-based *ROI*, which weights larger trades more heavily, reflects the activity of larger retail investors. It is reasonable to assume that smaller retail investors are more likely to rely on ChatGPT for trading guidance, so the outage effect should be more visible in the trade-based measure. Therefore, we use trade-based *ROI* as our primary proxy in the following tests. [13] We first test whether outages affect the magnitude of one-sidedness using the absolute value of signed *ROI*:[14]

$$Abs\_ROI_{i,t,p} = \alpha + \beta_1 Outage_{t,p} + \gamma Controls_{i,t} + \varepsilon_{i,t,p} \quad (4)$$

Table 7 shows no statistically significant change in *Abs_ROI* during outages (t= 1.36), suggesting that ChatGPT outages do not systematically affect the extent of one-sidedness of retail order.

[Insert Table 7 here]

We then examine whether ChatGPT affects the dynamics of net retail order flow. Following Bogousslavsky and Collin-Dufresne (2023), we define retail order imbalance volatility (*ROIV*) as the standard deviation of signed 5-minute ROI across all intervals within the event window for a given stock. We estimate:

$$ROIV_{i,t} = \alpha + \beta_1 Outage_t + \gamma Controls_{i,t} + \varepsilon_{i,t} \quad (5)$$

Table 7 shows that *ROIV* increases significantly during outages (0.006, t = 2.92). Since *Outage* captures periods when ChatGPT is unavailable, this implies that access to ChatGPT stabilizes net retail order flow. Collectively, these findings support the view that ChatGPT coordinates retail trading by reducing idiosyncratic fluctuations in investor beliefs.

4.3 The Effect of ChatGPT Outages on Retail Trading Quality

The documented increase in ChatGPT-driven retail synchronicity raises a fundamental question for market efficiency: does this reflect a collective improvement in information aggregation or an overreaction? To distinguish between these two possibilities, we examine whether retail *ROI* predicts subsequent returns and whether this predictability changes with ChatGPT availability, following prior work (e.g., Barber et al., 2022; Farrell et al., 2022;

[13] We also use volume-based ROI for all tests. As expected, the findings are qualitatively consistent but generally statistically weaker than those based on trade-based ROI, which is consistent with our assumption that smaller retail investors being more affected by ChatGPT outages.
[14] We focus on absolute ROI because our interest lies in the intensity of one-sidedness, not its direction.

Havakhor et al., 2025). If ChatGPT facilitates price discovery, *ROI* should effectively and persistently predict returns; If instead we observe short-term predictability followed by reversal, this would indicate that ChatGPT induces overreaction.

4.3.1 Return Predictability of Retail Order Imbalance

To allow for asymmetric effects on the buy and sell sides, we decompose signed *ROI* into selling *ROI* and buying *ROI*. Specifically, selling *ROI* retains the original value when *ROI* is negative and is set to zero otherwise; buying *ROI* retains the original value when *ROI* is positive and is set to zero otherwise (Chordia et al., 2002). We estimate return predictability at both intraday horizons (5–25 minutes) and multi-day horizons (1–5 days) using:

$$\begin{aligned} Ret_{i,t,[p+1,\ p+n]}\ (Ret_{i,[t+1,\ t+n]}) &= \alpha + \beta_1 SellROI_{i,t,p} * Outage_{t,p} \\ &+ \beta_2 BuyROI_{i,t,p} * Outage_{t,p} + \beta_3 SellROI_{i,t,p} \\ &+ \beta_4 BuyROI_{i,t,p} + \beta_5\, Outage_{t,p} + \gamma Controls_{i,t} + \varepsilon_{i,t,p} \end{aligned} \tag{6}$$

The results in Table 8 reveal a clear return reversal pattern on the sell side during normal periods, captured by $\beta 3$. Selling *ROI* significantly predicts returns at short intraday horizons (0.009, t = 3.61 at 5 minutes; 0.012, t = 2.90 at 10 minutes; 0.010, t = 2.08 at 15 minutes), but its coefficient loses significance at longer horizons and eventually reverses sign (−0.196, t = −2.44 at 3 days; −0.191, t = −1.99 at 5 days). This pattern is consistent with the behavioral hypothesis of investor overreaction (De Bondt and Thaler, 1985) rather than permanent price discovery (Barber et al., 2008; Barber et al., 2022). The interaction term *SellROI* × *Outage* is generally opposite in sign to the baseline selling *ROI* coefficient $\beta 3$, indicating that this reversal pattern is attenuated when ChatGPT is unavailable.

[Insert Table 8 here]

On the buy side, we find no return reversal and no significant outage interaction. One possible explanation is that retail buying is largely driven by attention and search frictions, whereas selling is constrained to positions already held and is therefore more closely tied to how investors process and act on information (Barber and Odean, 2008; Barber et al., 2022). The sell side thus provides a cleaner setting for studying the effect of an information-processing tool like ChatGPT, and we focus on retail selling for the remainder of this section.

4.3.2 Return Predictability of Decomposed Selling ROI

Prior studies point out that return reversals may also stem from compensation for liquidity provision (e.g. Barber et.al, 2008; Barrot et.al, 2016; Barber et.al, 2022). To isolate the overreaction channel, we follow Boehmer et al. (2021) and decompose selling *ROI* into three components: *ROI_persistence* (a proxy for price pressure), *ROI_liquidity* (a proxy for liquidity provision), and *ROI_other* (residual component).[15] The components are estimated from the following panel regression:

$$SellROI_{i,t,p} = \alpha + \beta_1 ROI_{i,t,[p-5,p-1]} + \beta_2 Ret_{i,t,[p-5,p-1]} + \varepsilon_{i,\mathrm{t},p} \tag{7}$$

where $\widehat{ROI}_{\text{Persistence}} = \hat{\beta}_1 ROI_{i,t,[p-5,p-1]}$, $\widehat{ROI}_{\text{liqudity}} = \hat{\beta}_2 Ret_{i,t,[p-5,p-1]}$ and $\widehat{ROI}_{\text{other}} = ROI_{i,t,p} - \widehat{ROI}_{\text{Persistence}} - \widehat{ROI}_{\text{liqudity}}$.

We then re-estimate the return predictability regressions, interacting each component with the outage indicator:

$$\begin{aligned} Ret_{i,t,[p+1,\ p+n]}\left(Ret_{i,[t+1,\ t+n]}\right) &= \alpha + \beta_1 ROI_{Other_{i,t,p}} * Outage_{t,p} + \beta_2 ROI_{Liquidity_{i,t,p}} * Outage_{t,p} \\ &+ \beta_3 ROI_{Persistence_{i,t,p}} * Outage_{t,p} + \gamma Controls_{i,t} + \varepsilon_{i,\mathrm{t},p} \end{aligned} \tag{8}$$

Table 9 shows that the reversal pattern and its outage-related attenuation are concentrated in *ROI_other* rather than *ROI_liquidity*. This indicates that the primary channel through which ChatGPT degrades retail trading quality is investor overreaction, rather than the compensation for liquidity provision.

[Insert Table 9 here]

4.3.3 Outage effects and information load

To further enhance the link between overreaction to ChatGPT-driven retail synchronicity, we examine whether the outage effect on *ROI* predictability weakens when more firm-specific information is publicly available. The intuition is straightforward: ChatGPT is less likely to generate homogeneous trading suggestions across users with abundant available information, the overreaction channel should be dampened, and the outage effect should therefore be smaller. We measure the volume of available information (*Information Load*) using the

[15] Farrell et.al. (2022) has applied decomposition method in high-frequency level.

number of firm-specific events recorded in Capital IQ Key Developments within the prior 30 days (Cao et al., 2024) and estimate:[16]

$$Ret_{i,t,[p+1,\ p+n]}\ \left(Ret_{i,[t+1,\ t+n]}\right)$$
$$= \alpha + \beta_1 SellROI_{i,t,p}\left(ROI_{Other_{i,t,p}}\right) * \ Outage_{t,p} \ * \ Information_{Load_{i,t}}$$
$$+ \gamma Controls_{i,t} + \varepsilon_{i,\mathrm{t},p} \quad (9)$$

Consistent with this prediction, Table 10 finds that higher information load attenuates the outage effect on retail trading quality, as reflected in the sign of the three-way interaction term being opposite to that of the baseline *ROI* (*ROI_other*) × *Outage* interaction.

[Insert Table 10 here]

Taken together, the findings in this section establish that the synchronized retail selling facilitated by ChatGPT reflects an aggregated overreaction to homogenized signals rather than informative price discovery. ChatGPT appears to amplify collective trading responses beyond what can be justified by fundamentals, which lowers average retail trading quality and introduces a new source of market fragility.

4.4 The Effect of ChatGPT Outages on Market Quality

Finally, we examine the market-wide consequences of ChatGPT-induced excessive synchronized retail trading. Previous research suggests that excessive correlated trading can create significant inventory risks for market makers, harming liquidity and raising volatility (e.g., Chordia et al., 2002; Park and Sabourian, 2011; Eaton et al., 2022).

4.4.1 Market Liquidity and the Contrarian Mechanism

We first test whether market liquidity changes during ChatGPT outages by estimating:

$$ILLiq_{i,t,p} = \alpha + \beta_1 Outage_{t,p} \ + \gamma Controls_{i,t} + \varepsilon_{i,\mathrm{t},p} \quad (10)$$

where $ILLiq$ is one of three illiquidity measures: price impact, quoted spread, and effective spread.

[16] In untabulated tests, we also measure firm information load using the number of firm-specific events over 90 and 180 days. The findings are held.

As reported in Table 11, we document a significant improvement in liquidity during ChatGPT outages, consistent with the findings in Chen et.al (2025) and Ecker et.al (2026). Specifically, price impact declines by 0.003, a 15% reduction relative to the sample mean. Quoted and effective spreads fall by 2.5% and 4.4%, respectively. These findings indicate that ChatGPT availability normally imposes a substantial liquidity cost on the market.

[Insert Table 11 here]

To understand the mechanism behind this liquidity improvement during the outage, we examine whether ChatGPT shifts retail investors away from their natural role as liquidity providers (Kaniel et al., 2008; Kelley and Tetlock, 2013). Prior work shows that investors provide liquidity when they trade contrarian to recent price moves, and consume liquidity when they chase trends (e.g., Kelley and Tetlock, 2013; Barrot et al., 2016; Eaton et al., 2022). Building on this, we regress retail order imbalances on recent returns (overnight return, one-day lagged return, and cumulative returns from day *t*−5 through *t*−2) for the full sample, buy and sell side separately:

$$ROI_{i,t,p} = \alpha + \beta_1 Outage_{t,p} * Ret_{i,t,overnight} + \beta_2 Outage_{t,p} * Ret_{i,t-1} + \beta_3 Outage_{t,p} * Ret_{i,[t-2,t-5]} + \gamma Controls_{i,t} + \varepsilon_{i,\mathrm{t},p} \tag{11}$$

Table 12 shows that during control periods, retail investors exhibit a modest but consistent contrarian bias relative to overnight returns (Full Sample: −0.003, t = −4.81; Buy-side: −0.001, t = −4.91; Sell-side: −0.002, t = −3.97). Crucially, this contrarian behavior strengthens significantly during outages across all specifications (Full Sample: −0.014, t = −3.04; Buy-side: −0.005, t = −3.12; Sell-side: −0.009, t = −2.86). This suggests that when retail investors lose access to ChatGPT, they revert to contrarian strategies more that supply liquidity to the market. Conversely, the availability of ChatGPT appears to shift retail investors away from their natural liquidity-providing role, likely by synchronizing retail attention toward common informational shocks.

[Insert Table 12 here]

4.4.2 Intraday and Systematic Volatility

We next examine whether ChatGPT outages influence stock return volatility by estimating:

$$Vola_{i,t,p} = \alpha + \beta_1 Outage_{t,p} + \gamma Controls_{i,t,\mathrm{p}} + \varepsilon_{i,\mathrm{t},p} \tag{12}$$

where *Vola* is the standard deviation of trade-by-trade returns within each 5-minute interval, computed for intervals with at least 10 trades.

Panel A of Table 13 shows that volatility declines by 0.298 during outages (t = −2.18), relative to a sample mean of 3.29, which corresponds to about a 9% reduction. Thus, ChatGPT availability is associated with higher intraday price fluctuations, consistent with our earlier finding that ChatGPT induces overreactions, which in turn create larger intraday price pressure.

[Insert Table 13 here]

To further isolate the source of this volatility, we follow Campbell et al. (2001) and decompose volatility into firm-specific idiosyncratic, industry-level systematic, and market-level systematic components. Specifically, for each 5-minute interval q in day t, we compute the value-weighted market return $r_{m,t,q} = \sum_i w_{i,t,q}\, r_{i,t,q}$, where $r_{j,t,q}$is the 5-minute return of stock i and $w_{i,t,q}$ is its market-capitalization weight. For each industry j, we compute the value-weighted industry return $r_{j,t,q} = \sum_{i \in j} w_{i,t,q}^{(j)}\, r_{i,t,q}$, where $w_{i,t,q}^{(j)}$ is stock i's weight in industry j. The 5-minute return of stock i in industry j is then decomposed as

$$r_{i,t,q} = r_{m,t,q} + \tilde{\eta}_{j,t,q} + \tilde{\varepsilon}_{i,t,q},$$

where the industry residual is $\tilde{\eta}_{j,t,q} = r_{j,t,q} - r_{m,t,q}$ and the firm-specific residual is $\tilde{\varepsilon}_{i,t,q} = r_{i,t,q} - r_{j,t,q}$.

We measure intraday volatilities using realized variance by summing squared components across 5-minute intervals within day t for each stock i:

$$RV_{i,t}^{MKT} = \sum_q r_{m,t,q}^2,\ RV_{i,t}^{IND} = \sum_q \tilde{\eta}_{j,t,q}^2,\ RV_{i,t}^{FIRM} = \sum_q \tilde{\varepsilon}_{i,t,q}^2.$$

The corresponding realized volatilities are $VOL_{i,t}^k = \sqrt{RV_{i,t}^k}$ for $k \in \{MKT, IND, FIRM\}$.

Panel B of Table 13 reports the results of estimating Equation (12) separately for each volatility component. All three components decline significantly during outages (firm-specific: −0.087, t = −4.02; industry-level: −0.037, t = −4.79; market-level: −0.066, t = −2.92). Importantly, the proportional decline is most pronounced at the systematic level:

while firm-specific volatility falls by 12.1%, industry- and market-level volatilities drop by 20.6% and 27.5%, respectively. This disproportionate impact on systematic component can be explained by Lee et.al (1991) noise-trader model: when retail noise becomes correlated across investors and across assets, the resulting noise-trader risk becomes non-diversifiable and turns into a systematic risk. Our decomposition therefore suggests that ChatGPT affects volatility not merely by amplifying individual-level behavioral biases, but primarily by synchronizing idiosyncratic noise, thereby increasing systematic risk.

Taken together, our findings provide early empirical evidence that ChatGPT may weaken market stability by synchronizing retail noise and amplifying the aggregate market impact of retail investors' behavioral biases. The results support the growing regulatory concern that widespread GenAI adoption may introduce an emerging channel of systematic risk in financial markets.

5 Conclusion

In this paper, we exploit unexpected full ChatGPT outages in 2023–2024 to provide early evidence on whether widespread GenAI adoption alters retail investors' behavioral biases and, through that channel, affects market quality. We show that ChatGPT availability synchronizes retail beliefs, increasing correlation in retail order flow and aggregating idiosyncratic retail noise into a coordinated market force with implications for systematic risk.

Our empirical analysis yields three main findings. First, ChatGPT availability increases the synchronicity of retail beliefs both within stocks across investors and across different stocks, confirming the role of ChatGPT as a "belief homogenizer." Second, this synchronization reflects aggregated overreaction rather than improved information aggregation: synchronized retail selling generates short-term price pressure followed by reversal, and the effect is concentrated in the behavioral component of order flow rather than in compensation for liquidity provision. Third, this synchronization imposes substantial market-wide costs. Liquidity is worse when ChatGPT is available, with higher price impact and wider spreads, and we trace this to the crowding out of retail investors' natural contrarian, liquidity-providing role. Intraday volatility is also higher, and the increase is disproportionately concentrated in systematic components. Our results suggest that ChatGPT affects market stability mainly by aggregating dispersed retail noise into undiversifiable systematic risk.

These findings highlight a fundamental tension in AI-assisted investing. At the individual level, ChatGPT lowers information costs and raises retail participation. At the collective level, it synchronizes beliefs, aggregates overreaction, crowds out liquidity provision, and elevates systematic risk. The net welfare effect depends on which force dominates, and our evidence suggests that the collective harm is substantial.

The policy implications are direct. Regulators have raised concerns that GenAI-induced herding could become a source of systemic vulnerability, but direct empirical evidence has been lacking. Our results provide evidence consistent with these concerns: a single dominant AI platform already generates measurable market-wide risk. As GenAI tools grow in sophistication and retail penetration, and as more trading decisions are delegated to AI-powered systems, the scale of these effects is likely to grow.

Several limitations are worth noting. Our analysis focuses on a single platform and a specific period. If investors substituted toward other GenAI tools during ChatGPT outages, our estimates likely understate the broader effect of GenAI availability. Besides, we cannot observe the content of ChatGPT's outputs or query-level usage during our sample period; future work with access to prompt-and-response data could illuminate the specific informational channels through which belief homogenization operates. We also do not observe individual investors' account-level trades, which prevents a direct link between an individual's GenAI usage and their synchronized trading or overreaction. Finally, our sample emphasizes stocks with meaningful retail activity, so effects in less retail-dominated market segments may differ.

Reference

Banerjee, A. V. (1992). A simple model of herd behavior. *The quarterly journal of economics*, *107*(3), 797-817.

Bank for International Settlements. (2024). *Artificial intelligence and the economy: Implications for central banks* (Chapter III, *Annual Economic Report 2024*). https://www.bis.org/publ/arpdf/ar2024e3.htm

Bank of England. (2025). *Financial stability in focus: Artificial intelligence in the financial system*. https://www.bankofengland.co.uk/financial-stability-in-focus/2025/april-2025

Barber, B. M., Huang, X., Odean, T., & Schwarz, C. (2022). Attention-induced trading and returns: Evidence from Robinhood users. *The Journal of Finance*, *77*(6), 3141-3190.

Barber, B. M., Huang, X., Jorion, P., Odean, T., & Schwarz, C. (2024a). A (sub) penny for your thoughts: Tracking retail investor activity in TAQ. *The Journal of Finance*, *79*(4), 2403-2427.

Barber, B. M., Lin, S., & Odean, T. (2024b). Resolving a paradox: Retail trades positively predict returns but are not profitable. *Journal of Financial and Quantitative Analysis*, *59*(6), 2547-2581.

Barber, B. M., & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The journal of Finance*, *55*(2), 773-806.

Barber, B. M., & Odean, T. (2002). Online investors: do the slow die first?. *The Review of financial studies*, *15*(2), 455-488.

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

Barber, B. M., Odean, T., & Zhu, N. (2008). Do retail trades move markets? *The Review of Financial Studies, 22*(1), 151-186.

Barrot, J. N., Kaniel, R., & Sraer, D. (2016). Are retail traders compensated for providing liquidity?. *Journal of Financial Economics*, *120*(1), 146-168.

Bertomeu, J., Lin, Y., Liu, Y., & Ni, Z. (2025). The impact of generative AI on information processing: Evidence from the ban of ChatGPT in Italy. *Journal of Accounting and Economics*, 101782.

Boehmer, E., Jones, C. M., Zhang, X., & Zhang, X. (2021). Tracking retail investor activity. *The Journal of Finance*, *76*(5), 2249-2305.

Bogousslavsky, V., & COLLIN-DUFRESNE, P. I. E. R. R. E. (2023). Liquidity, volume, and order imbalance volatility. *The Journal of Finance*, *78*(4), 2189-2232.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, *100*(5), 992-1026.

Blankespoor, E., Croom, J., & Grant, S. M. (2024). Generative AI and investor processing of financial information. *Available at SSRN 5053905*.

Blankespoor, E., deHaan, E., & Marinovic, I. (2020). Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics*, *70*(2-3), 101344.

Bradshaw, M. T., Ma, C., Yost, B., & Zou, Y. (2025). Generative AI use by capital market information intermediaries: Evidence from Seeking Alpha. *Available at SSRN*.

Bradley, D., Hanousek Jr, J., Jame, R., & Xiao, Z. (2024). Place your bets? The value of investment research on Reddit's Wallstreetbets. *The Review of Financial Studies*, *37*(5), 1409-1459.

Breitung, C., & Müller, S. (2025). Global business networks. *Journal of Financial Economics*, *166*, 104007.

Campbell, J. Y., Lettau, M., Malkiel, B. G., & Xu, Y. (2001). Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *The journal of finance*, *56*(1), 1-43.

Cao, S., Jiang, W., Wang, J., & Yang, B. (2024). From man vs. machine to man+ machine: The art and AI of stock analyses. *Journal of Financial Economics*, *160*, 103910.

Chang, A., Dong, X., Martin, X., & Zhou, C. (2023). AI democratization, return predictability, and trading inequality. *Available at SSRN 4543999*.

Chatterji, A., Cunningham, T., Deming, D. J., Hitzig, Z., Ong, C., Shan, C. Y., & Wadman, K. (2025). *How people use chatgpt* (No. w34255). National Bureau of Economic Research.

Cheng, Q., Lin, P., & Zhao, Y. (2025). Does generative AI facilitate investor trading? Early evidence from ChatGPT outages. *Journal of Accounting and Economics*, 101821.

Chordia, T., Roll, R., & Subrahmanyam, A. (2002). Order imbalance, liquidity, and market returns. *Journal of Financial economics*, *65*(1), 111-130.

Croom, J. (2025). Interactivity and illusions of ability: How using generative AI affects investor judgments. *Journal of Accounting Research*.

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under-and overreactions. *the Journal of Finance*, *53*(6), 1839-1885.

De Bondt, W. F., & Thaler, R. (1985). Does the stock market overreact?. *The Journal of finance*, *40*(3), 793-805.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of political Economy*, *98*(4), 703-738.

Deloitte Report (2024). Retail investors may soon rely on generative AI for financial advice. https://www2.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-predictions/2024/ai-financial-advisor-for-retail-investment.html?id=us:2el:3dp:wsjspon:awa:WSJCIO:2024:WSJFY25

Devenow, A., & Welch, I. (1996). Rational herding in financial economics. *European economic review*, *40*(3-5), 603-615.

Drake, M. S., Thornock, J. R., & Twedt, B. J. (2017). The internet as an information intermediary. *Review of Accounting Studies*, *22*(2), 543-576.

Eaton, G. W., Green, T. C., Roseman, B. S., & Wu, Y. (2022). Retail trader sophistication and stock market quality: Evidence from brokerage outages. *Journal of Financial Economics*, *146*(2), 502-528.

Ecker, F., Li, X., Li, Y., & Wu, F. (2026). How stock market participants use generative artificial intelligence: Evidence from user–platform interaction data. *Journal of Accounting Research, Forthcoming*.

European Securities and Markets Authority. (2023). *The financial stability implications of artificial intelligence*. https://www.esma.europa.eu/sites/default/files/library/ESMA50-164-6247-AI_in_securities_markets.pdf

eToro. (2023, July 11). Four in ten U.S. retail investors use or are open to using ChatGPT-style AI to help pick and manage investments https://www.etoro.com/en-us/news-and-analysis/latest-news/press-release/four-in-ten-u-s-retail-investors-use-or-are-open-to-using-chat-gpt-style-ai-to-help-pick-and-manage-investments/

Even-Tov, O., Lourie, B., & Nekrasov, A. (2025). The Effect of AI on Retail Investor Behavior: Evidence from Account-Level Trading Data. *Available at SSRN 5778246*.

Farrell, M., Green, T. C., Jame, R., & Markov, S. (2022). The democratization of investment research and the informativeness of retail investor trading. *Journal of Financial Economics*, *145*(2), 616-641.

Fisher, A., Martineau, C., & Sheng, J. (2022). Macroeconomic attention and announcement risk premia. *The Review of Financial Studies*, *35*(11), 5057-5093.

Financial Stability Board. (2024). *The financial stability implications of artificial intelligence*. https://www.fsb.org/uploads/P14112024.pdf

Forbes, “ChatGPT Hits 1 Billion Users? ‘Doubled In Just Weeks’ Says OpenAI CEO,” Forbes, April 12 2025.

Froot, K. A., Scharfstein, D. S., & Stein, J. C. (1992). Herd on the street: Informational inefficiencies in a market with short-term speculation. *The Journal of finance*, *47*(4), 1461-1484.

Garfinkel, J. A. (2009). Measuring investors' opinion divergence. *Journal of Accounting Research*, *47*(5), 1317-1348.

Graham, J. R. (1999). Herding among investment newsletters: Theory and evidence. *The Journal of Finance*, *54*(1), 237-268.

Hansen, A. L., & Lee, S. J. (2025). Financial Stability Implications of Generative AI: Taming the Animal Spirits. *arXiv preprint arXiv:2510.01451*.

Hasbrouck, J., & Seppi, D. J. (2001). Common factors in prices, order flows, and liquidity. *Journal of financial Economics*, *59*(3), 383-411.

Havakhor, T., Rahman, M. S., Zhang, T., & Zhu, C. (2025). Tech-enabled financial data access, retail investors, and gambling-like behavior in the stock market. Management Science, 71(2), 1646-1670.

Hillert, A., Jacobs, H., & Müller, S. (2014). Media makes momentum. *The Review of Financial Studies*, *27*(12), 3467-3501.

Holzman, E. R., Marshall, N. T., Schroeder, J. H., & Yohn, T. L. (2021). Is all disaggregation good for investors? Evidence from earnings announcements: E. Holzman et al. *Review of Accounting Studies*, *26*(2), 520-558.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of finance*, *54*(6), 2143-2184.

International Monetary Fund. (2024). *GLOBAL FINANCIAL STABILITY REPORT: Steadying the Course: Uncertainty, Artificial Intelligence, and Financial Stability*. https://www.imf.org/-/media/files/publications/gfsr/2024/october/english/textrevised.pdf

Kaniel, R., Saar, G., & Titman, S. (2008). Individual investor trading and stock returns. *The Journal of finance*, *63*(1), 273-310.

Kelley, E. K., & Tetlock, P. C. (2013). How wise are crowds? Insights from retail orders and stock returns. *The Journal of Finance*, *68*(3), 1229-1265.

Kremer, S., & Nautz, D. (2013). Causes and consequences of short-term institutional herding. *Journal of Banking & Finance*, *37*(5), 1676-1686.

Kumar, A., & Lee, C. M. (2006). Retail investor sentiment and return comovements. *The Journal of Finance*, *61*(5), 2451-2486.

Lee, C. M. (1992). Earnings news and small traders: An intraday analysis. *Journal of Accounting and Economics*, *15*(2-3), 265-302.

Lee, C. M., Shleifer, A., & Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle. *The journal of finance*, *46*(1), 75-109.

Liu, H., Peng, C., Xiong, W. A., & Xiong, W. (2022). Taming the bias zoo. *Journal of Financial Economics*, *143*(2), 716-741.

Liu, Yan; Huang, Jingyun; Wang, He. (2025). Who on Earth Is Using Generative AI? Global Trends and Shifts in 2025. Policy Research Working Paper; 11231. © World Bank. http://hdl.handle.net/10986/43859 License: CC BY 3.0 IGO.

Lopez-Lira, A., & Tang, Y. (2023). Can chatgpt forecast stock price movements? return predictability and large language models. *arXiv preprint arXiv:2304.07619*.

Lu, H., Wang, K. Q., & Wang, X. (2014). Price shocks, news disclosures, and asymmetric drifts. *The Accounting Review*, *89*(5), 1805-1834.

Lucca, D. O., & Moench, E. (2015). The pre-FOMC announcement drift. *The Journal of finance*, *70*(1), 329-371.

Odean, T. (1998). Volume, volatility, price, and profit when all traders are above average. *The journal of finance*, *53*(6), 1887-1934.

Oliver Wyman (2024). How generative AI is transforming business and society: The good, the bad, and everything in between. https://www.oliverwymanforum.com/content/dam/oliver-wyman/ow-forum/gcs/2023/AI-Report-2024-Davos.pdf

Park, A., & Sabourian, H. (2011). Herding and contrarian behavior in financial markets. *Econometrica*, *79*(4), 973-1026.

Piotroski, J. D., & Roulstone, D. T. (2004). The influence of analysts, institutional investors, and insiders on the incorporation of market, industry, and firm-specific information into stock prices. *The accounting review*, *79*(4), 1119-1151.

Reuters, “ChatGPT sets record for fastest-growing user base - analyst note,” Reuters, February 2 2023.

Reuters, “OpenAI hits $12 billion in annualized revenue, The Information reports,” Reuters, July 30 2025.

Savor, P., & Wilson, M. (2013). How much do investors care about macroeconomic risk? Evidence from scheduled economic announcements. *Journal of Financial and Quantitative Analysis*, *48*(2), 343-375.

Scharfstein, D. S., & Stein, J. C. (1990). Herd behavior and investment. *The American economic review*, 465-479.

Shleifer, A., & Summers, L. H. (1990). The noise trader approach to finance. *Journal of Economic perspectives*, *4*(2), 19-33.

Shiller, R. 2006. Irrational exuberance. Danvers, MA: Crown Business Publishing.

Trueman, B. (1994). Analyst forecasts and herding behavior. *The review of financial studies*, *7*(1), 97-124.

Watkins, J. (2022). Consequences of prescribed disclosure timeliness: Evidence from acceleration of the Form 8-K filing deadline. *The Accounting Review*, *97*(7), 429-463.

Tables

Table 1: Sample Selection

	Observations	Securities
Initial Sample:		
Intraday 5-minute intervals for U.S. ordinary common stocks listed on NYSE, NYSE American, or Nasdaq (excluding non-ordinary shares and sub-$1 stocks) during documented ChatGPT outage windows, plus matched control intervals (the same 5-minute intervals (same time-of-day bins) for the same firm over the prior five trading days), from 2023 to 2024.	5,682,191	4,201
Less:		
Observations with fewer than 500 average daily retail trades during the control period.	(3,448,427)	(1,155)
Intervals with fewer than 10 retail trades	(525,460)	(354)
Observations with missing variables	(146,611)	(352)
Final Sample	**1,561,693**	**2,340**

This table reports the sample selection process. The final sample for analyses includes 1,561,693 security-date-5-minute interval observations from 2,340 stock securities from 2023 to 2024.

Table 2: Summary Statistics

Variable	N	Mean	SD	p10	p50	p75	p90
Abs Trade-based ROI (%)	1,561,693	29.36	21.97	4.72	25.00	42.86	62.50
Abs Volume-based ROI (%)	1,561,693	41.45	27.68	6.98	37.60	62.51	83.33
BEME	1,561,693	0.47	0.46	0.08	0.34	0.64	1.00
ΔTO	1,554,520	0.02	1.04	-0.12	0.01	0.09	0.21
Effective Spread	1,561,693	0.09	0.11	0.01	0.06	0.11	0.19
Idiosyncratic Volatility	104,368	0.72	0.85	0.12	0.45	0.86	1.57
Log (Intraday Volatility)	1,561,693	-11.20	1.37	-12.83	-11.37	-10.37	-9.39
Log (ME)	1,561,693	9.29	1.72	7.04	9.37	10.54	11.46
Log (Number of Analysts)	1,561,693	2.51	0.58	1.79	2.64	2.94	3.14
Log (Price)	1,561,693	3.99	1.20	2.35	4.09	4.82	5.46
MATO	1,561,709	-1.16	1.57	-1.51	-1.17	-1.03	-0.89
Price Impact	1,561,693	0.02	0.17	-0.12	0.01	0.06	0.17
SYNCH (Within-industry)	96,760	-1.39	2.45	-4.34	-1.12	0.05	1.13
SYNCH (Market)	97,021	-1.43	2.40	-4.31	-1.14	-0.07	0.98
SYNCH (Overall)	89,149	-0.43	1.91	-2.56	-0.50	0.51	1.68
Quoted Spread	1,561,693	0.12	0.14	0.03	0.08	0.14	0.25
ROIV (Trade-based)	100,707	0.45	0.17	0.25	0.45	0.55	0.66
ROIV (Volume-based)	100,707	0.31	0.13	0.15	0.29	0.38	0.47
Retail Trading Count	1,561,693	65.20	93.78	15.00	36.00	68.00	132.00
Retail Trading Volume	1,561,693	0.79	1.96	0.05	0.22	0.58	1.60
Return [-1] (%)	1,561,693	0.05	5.55	-3.26	-0.02	1.19	2.86
Return [-5, -2] (%)	1,561,693	-0.44	10.27	-7.39	-0.49	2.11	5.62
Return Overnight (%)	1,561,693	-0.01	1.50	-1.46	-0.02	0.49	1.31
Signed Trade-based ROI (%)	1,561,693	0.01	0.37	-0.48	0.01	0.26	0.48
Signed Volume-based ROI (%)	1,561,693	-0.03	0.50	-0.71	-0.03	0.35	0.66
SUV	1,558,524	0.01	1.24	-0.87	-0.17	0.31	1.00

Systematic Volatility (Industry)	104,368	0.18	0.15	0.04	0.12	0.24	0.40
Systematic Volatility (Market)	104,368	0.24	0.19	0.06	0.19	0.33	0.48
Volatility (bps)	1,561,693	3.29	3.75	0.69	2.13	3.82	6.88

This table reports the descriptive statistics of the main variables used in empirical analyses. Variables are measured at the stock–date–interval level unless otherwise noted. We winsorize continuous variables at the 1st and 99th percentiles, except stock returns. Sample sizes vary across variables because some measures are defined at higher aggregation levels or require additional data availability. See Appendix B for detailed variable definitions.

Table 3: ChatGPT Outages and Retail Trading Intensity

	Retail Trading Volume	Retail Trading Number
Outage	-0.044**	-4.100**
	(-2.30)	(-2.40)
Controls	Y	Y
Firm FE	Y	Y
Event Index × Time Interval FE	Y	Y
Obs	1,561,693	1,561,693
Adjusted R2	0.589	0.648

This table reports OLS regression results on the effect of ChatGPT outages on retail trading intensity measured at 5-minute intervals. The dependent variables are *Retail Trading Volume* and *Retail Trade Count*. The key independent variable is *Outage*, an indicator equal to one if the 5-minute interval overlaps a ChatGPT outage window and zero otherwise. Control variables include *Return Overnight, Return [-1], Return [-5, -2], Log (Intraday Volatility), Log (ME), BEME, Log (Price), and Log (Number of Analysts)*. Variable definitions are provided in Appendix B. All regressions include firm fixed effects and event-by-interval fixed effects. Standard errors are two-way clustered by firm and by event-by-date. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 4: Robustness Tests

	Retail Trading Volume	Retail Trading Number
Panel A. Exclude firm event days		
Outage	-0.044**	-4.238**
	(-2.30)	(-2.35)
Controls	Y	Y
Firm FE	Y	Y
Event Index × Time Interval FE	Y	Y
Obs	1,286,498	1,286,498
Adjusted R2	0.580	0.626
Panel B. Exclude Marco market events		
Outage	-0.043**	-3.912**
	(-2.02)	(-2.06)
Controls	Y	Y
Firm FE	Y	Y
Event Index × Time Interval FE	Y	Y
Obs	1,386,641	1,386,641
Adjusted R2	0.590	0.655
Panel C. Exclude outages less than 30 min		
Outage	-0.045**	-4.190**
	(-2.25)	(-2.34)
Controls	Y	Y
Firm FE	Y	Y
Event Index × Time Interval FE	Y	Y
Obs	1,493,382	1,493,382
Adjusted R2	0.590	0.647
Panel D. Exclude outages that begin before 10:00 AM		
Outage	-0.043**	-6.066**
	(-2.23)	(-2.53)
Controls	Y	Y
Firm FE	Y	Y
Event Index × Time Interval FE	Y	Y
Obs	1,008,421	1,008,421
Adjusted R2	0.591	0.682
Panel E. Pseudo outages		
Outage	-0.019	-2.358
	(-1.34)	(-1.64)
Controls	Y	Y
Firm FE	Y	Y
Event Index × Time Interval FE	Y	Y
Obs	1,561,693	1,561,693
Adjusted R2	0.589	0.647

This table reports robustness tests for the baseline specification in Table 3. Panel A excludes observations with firm-specific information events. Panel B excludes outage events that overlap major scheduled U.S. macro announcements (CPI, PPI, and FOMC Minutes). Panel C excludes outage events with durations shorter than 30 minutes. Panel D excludes outage events that begin before 10:00 AM ET. Panel E implements a placebo test by assigning the outage window to day t-1 (pseudo-outage). Control variables, fixed effects, clustering, and significance notation follow Table 3. Variable definitions are provided in Appendix B. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 5: ChatGPT Outages and Within-stock Retail Opinion Divergence

	MATO	ΔTO	SUV
Outage	0.107***	0.103***	0.064***
	(3.10)	(3.14)	(2.66)
Controls	Y	Y	Y
Firm FE	Y	Y	Y
Event Index × Time Interval FE	Y	Y	Y
Obs	1,561,693	1,554,512	1,558,508
Adjusted R2	0.335	0.054	0.027

This table reports OLS regression results on the effect of ChatGPT outages on retail investors' opinion divergence measured at 5-minute intervals. The dependent variables are *Market-adjusted Turnover (MATO), Abnormal Market-adjusted Turnover (ΔTO)*, and *Standarized Unexplained Volume (SUV)*. The key independent variable is *Outage*. Control variables follow the baseline specification (see Appendix B for definitions). All regressions include firm fixed effects and event-by-interval fixed effects, and standard errors are two-way clustered by firm and by event-by-date. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 6: ChatGPT Outages and Cross-stock Coordination

Panel A.OLS Regression

	SYNCH		
	(1)	(2)	(3)
Outage	-0.219*	-0.251*	-0.179*
	(-1.76)	(-1.96)	(-1.87)
Controls	Y	Y	Y
Firm FE	Y	Y	Y
Industry FE	Y	N	Y
Event Index FE	Y	Y	Y
Obs	96,649	96,908	89,025
Adjusted R2	0.197	0.154	0.306

Panel B. PCA

Percentage of Variance Explained by Principal Components

PC	Outage	Controls	Diff	t-value
Within-Industry				
PC1_Return	0.141	0.208	-0.067***	-11.423
PC1_ROI	0.015	0.025	-0.010***	-6.521
Market				
PC1	0.114	0.170	-0.056***	-3.091
PC1_ROI	0.030	0.032	-0.002	-0.731

This table reports evidence on the effect of ChatGPT outages on cross-stock coordination. Panel A reports OLS regressions of price synchronicity *(SYNCH=ln(R2/ (1-R2))*) on *Outage*, where for each firm-day *R2* is estimated from regressing the firm's 5-minute returns within the matched event window on the leave-one-out value-weighted Fama–French 12 industry return (Column 1), the value-weighted market return (Column 2), or both (Column 3); we require at least 10 firms per industry–interval for Columns (1) and (3). Control variables follow the baseline specification (see Appendix B for definitions). Columns (1) and (3) regression include firm fixed effects, industry fixed effects, and event fixed effects. Column (2) includes firm fixed effects and event fixed effects. Standard errors are two-way clustered by firm and by event-by-date. Panel B reports industry level and market level principal components analysis (PCA) results for 5-minute stock returns and ROI in outage windows and matched control windows; it reports the share of variation explained by the first principal component, the outage–control difference, as well as the t-value for differences. Appendix D describes the PCA procedure. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 7: ChatGPT Outages and Retail Order Imbalance and Volatility

	Abs_ROI	ROIV
Outage	0.467	0.006***
	(1.36)	(2.92)
Controls	Y	Y
Firm FE	Y	Y
Event Index FE	N	Y
Event Index × Time Interval FE	Y	N
Obs	1,561,693	100,606
Adjusted R2	0.115	0.326

This table reports OLS regression results on the effect of ChatGPT outages on retail order imbalance and retail order imbalance volatility. The dependent variables are *absolute retail order imbalance (Abs_ROI)* and *retail order imbalance volatility (ROIV)*. The key independent variable is *Outage*. Control variables follow the baseline specification (see Appendix B for definitions). The *Abs_ROI* regressions include firm fixed effects and event-by-interval fixed effects; the *ROIV* regressions include firm fixed effects and event fixed effects. Standard errors are both two-way clustered by firm and by event-by-date. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 8: ChatGPT Outages and the Predictability of Retail Order Imbalance

	Return t, [p+1]	Return t, [p+1, p+2]	Return t, [p+1, p+3]	Return t, [p+1, p+5]	Return [t+1]	Return [t+1, t+3]	Return [t+1, t+5]
Outage x Selling ROI	-0.013*	-0.022*	-0.030*	-0.041*	0.544*	0.484*	0.290
	(-1.81)	(-1.83)	(-1.87)	(-1.68)	(1.88)	(1.87)	(1.08)
Outage x Buying ROI	0.007	0.002	-0.002	0.001	-0.156	-0.545***	-0.284
	(1.21)	(0.18)	(-0.13)	(0.02)	(-0.73)	(-2.79)	(-1.19)
Selling ROI	0.009***	0.012***	0.010**	0.007	-0.092	-0.196**	-0.191**
	(3.67)	(2.90)	(2.08)	(0.94)	(-1.53)	(-2.44)	(-1.99)
Buying ROI	0.007***	0.012***	0.019***	0.019*	0.081	0.160	0.239*
	(2.68)	(2.81)	(2.97)	(1.79)	(1.14)	(1.50)	(1.92)
Outage	-0.005	-0.009	-0.016	-0.033	0.618*	-0.186	-0.086
	(-0.92)	(-0.87)	(-0.99)	(-1.24)	(1.81)	(-0.34)	(-0.14)
Controls	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y
Event Index FE× Time Interval FE	Y	Y	Y	Y	Y	Y	Y
Obs	1,549,585	1,494,659	1,406,620	1,238,014	1,561,653	1,561,653	1,561,653
Adjusted R2	0.022	0.027	0.031	0.039	0.106	0.205	0.242

This table reports OLS regression results on the relation between *retail order imbalance (ROI)* and subsequent returns, conditional on ChatGPT outages. The dependent variable is the stock's future return measured over intraday horizons (next 5–25 minutes) and multi-day horizons (1–5 days), as indicated by the column headers. The key independent variables are *SellROI* and *BuyROI* and their interactions with *Outage*, where *SellROI* retains the original value when *ROI* is negative and is set to zero otherwise; *BuyROI* retains the original value when *ROI* is positive and is set to zero otherwise; *Outage* is an indicator variable. Control variables follow the baseline specification (see Appendix B for definitions). All regressions include firm fixed effects and event-by-interval fixed effects, and standard errors are two-way clustered by firm and by event-by-date. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 9: ChatGPT Outages and the Predictability of Decomposed Selling ROI

	Return t, [p+1]	Return t, [p+1, p+2]	Return t, [p+1, p+3]	Return t, [p+1, p+5]	Return [t+1]	Return [t+1, t+3]	Return [t+1, t+5]
Outage x ROI_other	-0.011	-0.023*	-0.026*	-0.040*	0.517*	0.670***	0.486*
	(-1.49)	(-1.67)	(-1.68)	(-1.84)	(1.92)	(2.79)	(1.93)
Outage x ROI_liquidity	-0.004	-0.001	-0.012	-0.049	-0.009	-0.141	0.122
	(-0.32)	(-0.06)	(-0.31)	(-0.79)	(-0.05)	(-0.50)	(0.42)
Outage x ROI_persistence	-0.003	-0.013	-0.022	-0.034	1.057**	0.733	0.349
	(-0.26)	(-0.71)	(-0.86)	(-0.84)	(2.53)	(1.36)	(0.65)
ROI_other	0.010***	0.012***	0.012**	0.014*	-0.095	-0.160**	-0.142
	(3.56)	(2.72)	(2.11)	(1.75)	(-1.58)	(-2.10)	(-1.66)
ROI_liquidity	0.001	0.000	-0.006	-0.003	0.191	0.483**	0.564**
	(0.11)	(0.02)	(-0.22)	(-0.09)	(1.51)	(2.10)	(2.52)
ROI_persistence	0.005	0.007	0.001	-0.004	-0.001	-0.151	0.116
	(0.95)	(0.77)	(0.08)	(-0.19)	(-0.01)	(-0.91)	(0.56)
Outage	-0.004	-0.008	-0.011	-0.022	0.658**	-0.096	-0.063
	(-0.67)	(-0.76)	(-0.73)	(-0.93)	(2.02)	(-0.19)	(-0.11)
Controls	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y
Event Index × Time Interval FE	Y	Y	Y	Y	Y	Y	Y
Obs	702,293	670,956	628,122	546,211	702,278	702,278	702,278
Adjusted R2	0.03	0.034	0.039	0.049	0.106	0.215	0.259

This table reports OLS regression results on the relation between decomposed *selling ROI* components and subsequent returns, conditional on ChatGPT outages. The dependent variable is future return over the horizons shown in the column headers. The key independent variables are *ROI_persistence, ROI_liquidity*, and *ROI_other* and their interactions with *Outage*. These components are estimated as the fitted values from the panel regression:

$$SellROI_{i,t,p} = \alpha + \beta_1 ROI_{i,t,[p-5,p-1]} + \beta_2 Ret_{i,t,[p-5,p-1]} + \varepsilon_{i,\mathrm{t},p} ,$$

where $\widehat{ROI}_{\text{Persistence}} = \hat{\beta}_1 ROI_{i,t,[p-5,p-1]}$, $\widehat{ROI}_{\text{liqudity}} = \hat{\beta}_2 Ret_{i,t,[p-5,p-1]}$ and $\widehat{ROI}_{\text{other}} = ROI_{i,t,p} - \widehat{ROI}_{\text{Persistence}} - \widehat{ROI}_{\text{liqudity}}$ respectively. Control variables follow the baseline specification (see Appendix B for definitions). Control variables follow the baseline specification (see Appendix B for definitions). All regressions include firm fixed effects and event-by-interval fixed effects, and standard errors are two-way clustered by firm and by event-by-date. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 10: Heterogeneity by Firm-specific Information Load

	Return t, [p+1]	Return t, [p+1, p+2]	Return t, [p+1, p+3]	Return t, [p+1, p+5]	Return [t+1]	Return [t+1, t+3]	Return [t+1, t+5]
Panel A. Selling ROI							
Outage x ROI x Information_Load	0.012***	0.020**	0.023**	0.047**	-0.627***	-0.565**	-0.476*
	(3.27)	(2.45)	(2.07)	(2.43)	(-3.62)	(-2.38)	(-1.66)
Controls	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y
Event Index × Time Interval FE	Y	Y	Y	Y	Y	Y	Y
Obs	722,770	697,387	654,544	572,624	728,724	728,724	728,724
Adjusted R2	0.031	0.035	0.039	0.048	0.106	0.218	0.259
Panel B. Decomposed Selling ROI							
Outage x ROI_other x Information_Load	0.012***	0.015*	0.010	0.027**	-0.478***	-0.290*	-0.331
	(2.77)	(1.97)	(1.12)	(2.18)	(-4.41)	(-1.77)	(-1.59)
Controls	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y
Event Index × Time Interval FE	Y	Y	Y	Y	Y	Y	Y
Obs	702,293	670,956	628,122	546,211	702,278	702,278	702,278
Adjusted R2	0.03	0.035	0.039	0.049	0.106	0.215	0.26

This table reports OLS regression results testing whether the outage effect on ROI predictability varies with the firm's information environment. The dependent variable is future return over the horizons shown in the column headers. The key independent variable is the triple interaction *Outage* × *ROI* × *Information Load*, where *Information Load* is the number of firm-specific information events in Capital IQ Key Developments over the prior 30 days. Panel A uses *SellROI*; Panel B uses the decomposed *ROI_other* component. Control variables follow the baseline specification (see Appendix B for definitions). All regressions include firm fixed effects and event-by-interval fixed effects, and standard errors are two-way clustered by firm and by event-by-date. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 11: ChatGPT Outages and Illiquidity

	Price Impact	Quoted Spread	Effective Spread
Outage	-0.003**	-0.003**	-0.004**
	(-2.03)	(-2.50)	(-2.17)
Controls	Y	Y	Y
Firm FE	Y	Y	Y
Event Index × Time Interval FE	Y	Y	Y
Obs	1,561,693	1,561,693	1,561,693
Adjusted R2	0.023	0.842	0.722

This table reports OLS regression results on the effect of ChatGPT outages on market liquidity measured at 5-minute intervals. The dependent variables are *Price Impact, Quoted Spread*, and *Effective Spread*. The key independent variable is *Outage*. Control variables follow the baseline specification (see Appendix B for definitions). All regressions include firm fixed effects and event-by-interval fixed effects, and standard errors are two-way clustered by firm and by event-by-date. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 12: ChatGPT Outages and Liquidity Provision

	ROI		
	Full Sample	Buying	Selling
Outage x Return_Overnight	-0.014***	-0.005***	-0.009***
	(-3.04)	(-3.12)	(-2.86)
Outage x Return [-1]	-0.031	-0.007	-0.023
	(-0.67)	(-0.30)	(-0.84)
Outage x Return [-5, -2]	0.002	-0.002	0.004
	(0.09)	(-0.18)	(0.24)
Return_Overnight	-0.003***	-0.001***	-0.002***
	(-4.81)	(-4.91)	(-3.97)
Return [-1]	0.011	-0.012	0.022
	(0.46)	(-1.05)	(1.43)
Return [-5,-2]	-0.000	-0.002	0.001
	(-0.01)	(-0.17)	(0.22)
Outage	-0.006*	-0.001	-0.005*
	(-1.74)	(-0.35)	(-1.74)
Controls	Y	Y	Y
Firm FE	Y	Y	Y
Event Index × Time Interval FE	Y	Y	Y
Obs	1,561,693	1,561,693	1,561,693
Adjusted R2	0.052	0.056	0.081

This table reports OLS regression results testing whether retail order imbalance responds differently to recent returns during ChatGPT outages, reported for the full sample and separately for buy-side and sell-side ROI. The dependent variable is *ROI.* The key independent variables are interactions between *Outage* and recent returns (*Return Overnight, Return [-1], and Return [-5, -2]*). Control variables follow the baseline specification (see Appendix B for definitions). All regressions include firm fixed effects and event-by-interval fixed effects, and standard errors are two-way clustered by firm and by event-by-date. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 13: ChatGPT Outages and Volatility

Panel A. Volatility

	Volatility
Outage	-0.298**
	(-2.18)
Controls	Y
Firm FE	Y
Event Index × Time Interval FE	Y
Obs	1,561,693
Adjusted R2	0.533

Panel B. Idiosyncratic Volatility

		Systematic Volatility	
	Idiosyncratic Volatility	Industry	Market
Outage	-0.087***	-0.037***	-0.066***
	(-4.02)	(-4.79)	(-2.92)
Controls	Y	Y	Y
Firm FE	Y	Y	Y
Event Index FE	Y	Y	Y
Obs	104,286	104,286	104,286
Adjusted R2	0.577	0.738	0.69

This table reports OLS regression results on the effect of ChatGPT outages on return volatility. Panel A uses 5-minute *trade-based volatility* as the dependent variable. Panel B reports results for volatility components: *idiosyncratic volatility, industry systematic volatility,* and *market systematic volatility*. The decomposition method follows Campbell et al. (2001). Control variables follow the baseline specification (see Appendix B for definitions). Panel A includes firm fixed effects and event-by-interval fixed effects; Panel B includes firm fixed effects and event fixed effects, consistent with the component construction. Standard errors are both two-way clustered by firm and by event-by-date. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Appendices

Appendix A: ChatGPT Outage Events

This appendix lists the ChatGPT outage events in our sample. We collect outage dates, severity, and timing from OpenAI's public incident information. For outages that extend beyond regular trading hours (9:30–16:00 ET), we truncate the outage window to trading hours by setting the start time to 9:30 and/or the end time to 16:00. We aggregate trading data into 5-minute intervals and treat any interval that overlaps with the outage window as part of the outage period.

Date	Outage Severity	Starting Time (ET)	Ending Time (ET)	Duration (min)	Day of Week
01/18/2023	Full outage	12:04	13:18	74	Wed
01/31/2023	Full outage	12:30	13:34	64	Tues
02/14/2023	Full outage	12:59	14:37	98	Tues
02/21/2023	Full outage	09:30	16:00	390	Tues
02/27/2023	Full outage	10:00	13:30	210	Mon
03/15/2023	Full outage	13:55	14:18	23	Wed
03/20/2023	Full outage	12:41	16:00	199	Mon
05/09/2023	Full outage	15:00	15:19	19	Tues
05/16/2023	Full outage	14:46	15:21	35	Tues
05/24/2023	Full outage	15:44	16:00	16	Wed
06/30/2023	Full outage	14:41	15:20	39	Fri
08/29/2023	Full outage	10:04	11:16	72	Tues
08/31/2023	Full outage	09:52	14:03	251	Thu
09/13/2023	Full outage	09:30	11:01	91	Wed
11/08/2023	Full outage	09:30	10:46	76	Wed
03/18/2024	Full outage	14:52	15:48	56	Mon
05/21/2024	Full outage	14:40	15:48	68	Tues
06/04/2024	Full outage	10:33	13:17	164	Tues
08/21/2024	Full outage	12:27	13:02	35	Wed
10/17/2024	Full outage	13:31	13:51	20	Thu
12/26/2024	Full outage	14:00	16:00	120	Thu
Sum Duration				**2120**	
Avg Duration				**100.95**	

Appendix B: Variable Definitions

Variable	Definition
Book-to-Market (BEME)	The ratio of the book value of equity to the market value of equity.
Abnormal Market-adjusted Turnover (ΔTO)	Following Garfinkel (2009), ΔTO equals MATO minus its average level over the matched control window (prior trading days) for the same stock and time-of-day bin within the event.
Effective Spread	Equal-weighted average of the effective spread during the 5-min interval, then multiplied by 100. For each transaction, the effective spread is defined as $2 \times \lvert\ln(Pk) - \ln(Mk)\rvert$, where P is the trade price and M is the prevailing midquote.
Idiosyncratic Volatility	Following Campbell et.al (2001), intraday firm-specific (idiosyncratic) volatility, computed for each stock as the square root of the sum across 5-minute bins of $(r\_j - r\_ind)^2$, where r_t is the stock's interval return and r_ind is the value-weighted return of the stock's industry. We use Fama–French 12 industry classification.
Information Load	The logarithm of one plus the number of information events within 30 days. Number of information events, which is the number of firm-specific information events within 30 days in Capital IQ Key Development data and represents the volume of available information about the firm.
Log (Intraday Volatility)	The logarithm of trade-based intraday volatility multiplied by 100.
Log (Market Equity (ME))	The logarithm of market equity. ME = Price x Shrout, where SHROUTi,t is the total number of shares outstanding of stock i on day t.
Log (Number of Analysts)	The logarithm of one plus the number of analysts following the firm
Log (Price)	The logarithm of one plus the daily stock price
Market-adjusted Turnover (MATO)	Following Garfinkel (2009), MATO equals the stock's 5-minute retail turnover (retail share volume divided by shares outstanding) minus the market-wide turnover computed over the same date and 5-minute time-of-day bin.
Outage	An indicator variable that equals to one if the 5-minute interval overlaps with a ChatGPT outage window, and zero otherwise.

Price Impact	Equal-weighted average of the price impact of trades in a 5-min interval, then multiplied by 100. For each trade, the price impact is calculated as 2 × D × (Mt+5 – Mt)/Mt, where D equals 1 for a buy transaction and −1 for a sell transaction, Mt is the prevailing bid-ask mid-point, and Mt+5 is the bid-ask mid-point 5 min after the trade. We classify each transaction into a buy or sell transaction using Lee and Ready's (1991) approach.
Price Synchronicity (SYNCH)	For each firm-day, we compute R^2 from regressions of the firm's 5-minute returns (ROI) (within the matched event window) on (i) the leave-one-out value-weighted Fama–French 12 industry return (industry SYNCH), (ii) the value-weighted market return (market SYNCH), or (iii) both returns (overall systematic SYNCH), using the same 5-minute intervals. We require at least 10 firms per industry–interval to form the industry benchmark. We define SYNCH $= \ln(R^2/(1-R^2))$.
Quoted Spread	Equal-weighted average of the best bid-ask spread scaled by the mid-quote across all quotes in a 5-min interval, then multiplied by 100.
Retail Order Imbalance (ROI)	Signed imbalance scaled by total retail activity. Volume-based ROI is (RetailBuyVol−RetailSellVol) / (RetailBuyVol+RetailSellVol); trade-based ROI replaces volume with trade counts. We report ROI multiplied by 100.
Retail Order Imbalance Volatility (ROIV)	Following Bogousslavsky and COLLIN-DUFRESNE (2023), Retail Order Imbalance Volatility is the standard deviation of the five-minute (Trade-based or Volume-based) Retail Order Imbalance over the day.
Retail Trading Count	Number of retail trades in the 5-minute interval.
Retail Trading Volume	The total share volume traded by retail investors in a 5-min interval., then divided by 10,000.
Return [-1]	Stock return on day t-1.
Return [-5, -2]	Cumulative stock return from day t-5 through day t-2
Return Overnight	Stock return measured from the closing price on day t-1 to the opening price on day t,

Return t, [p+1, p+n]	The cumulative return over short intraday horizons n (5–25 minutes) For example, Return t, [p+1,p+3] represents the return for the next three 5-min intervals.
Return [t+1, t+n]	The cumulative return over longer daily horizons n (1–5 days). For example, Return[t+1,t+3] represents the return over the next three days.
Standardized Unexplained Volume (SUV)	Following Garfinkel (2009), for each stock and 5-minute interval, SUV is the residual from a control-period regression of retail turnover on absolute returns (allowing different slopes for positive vs. negative returns) and time-of-day effects, scaled by the standard deviation of the residuals in the control window.
Systematic Volatility (Industry)	Following Campbell et.al (2001), intraday industry-level residual (systematic) volatility, computed for each stock as the square root of the sum across 5-minute bins of $(r\_ind - r\_mkt)^2$, where r_ind is the value-weighted return of the stock's industry and r_mkt is the value-weighted market return. We use Fama–French 12 industry classification.
Systematic Volatility (Market)	Following Campbell et.al (2001), intraday market-level realized volatility, computed as the square root of the sum of squared value-weighted market returns across all 5-minute bins within the trading day.
Volatility	Standard deviation of trade-based returns in a 5-min interval, then multiplied by 10,000. We require more than ten trades for the 5-min interval for the calculation of this variable.

Appendix C: Supplementary Robustness Tests

Table C.1: Robustness Tests (A)

	ΔTO	Overall SYNCH	Abs_ROI	ROIV	Effective Spread	Volatility	Idiosyncratic Volatility	Industry-level Systematic Volatility
Panel A. Exclude firm event days								
Outage	0.099***	-0.190*	0.602	0.007***	-0.004**	-0.308**	-0.086***	-0.039***
	(3.03)	(-1.89)	(1.58)	(3.00)	(-2.17)	(-2.07)	(-3.72)	(-4.50)
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	N	Y	N	N	N	N	N	N
Event Index FE	N	Y	N	Y	N	N	N	N
Event Index × Time Interval FE	Y	N	Y	N	Y	Y	Y	Y
Obs	1,280,496	73,980	1,561,693	83,965	1,561,693	1,561,693	87,167	87,167
Adjusted R2	0.088	0.301	0.114	0.313	0.722	0.533	0.595	0.730
Panel B. Exclude Marco market events								
Outage	0.106***	-0.264***	0.496	0.006***	-0.005**	-0.318**	-0.094***	-0.044***
	(2.95)	(-2.78)	(1.29)	(2.70)	(-2.07)	(-2.13)	(-3.73)	(-4.91)
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	N	Y	N	N	N	N	N	N
Event Index FE	N	Y	N	Y	N	N	N	N
Event Index × Time Interval FE	Y	N	Y	N	Y	Y	Y	Y
Obs	1,379,786	77,089	1,386,641	83,248	1,386,641	1,386,641	86,219	86,219
Adjusted R2	0.061	0.302	0.119	0.333	0.717	0.527	0.582	0.736
Panel C. Exclude outages less than 30 min								

Outage	0.108***	-0.221**	0.423	0.004**	-0.004**	-0.289**	-0.097***	-0.044***
	(3.20)	(-2.02)	(1.19)	(2.12)	(-2.07)	(-2.04)	(-3.93)	(-5.27)
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	N	Y	N	N	N	N	N	N
Event Index FE	N	Y	N	Y	N	N	N	N
Event Index × Time Interval FE	Y	N	Y	N	Y	Y	Y	Y
Obs	1,486,421	71,999	1,493,382	79,427	1,493,382	1,493,382	86,429	86,429
Adjusted R2	0.054	0.229	0.114	0.467	0.726	0.539	0.570	0.721
Panel D. Exclude outages that begin before 10:00 AM								
Outage	0.109***	-0.154	0.872*	0.008***	-0.006*	-0.350	-0.083***	-0.031***
	(2.66)	(-1.32)	(1.72)	(3.25)	(-1.73)	(-1.60)	(-3.41)	(-3.22)
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	N	Y	N	N	N	N	N	N
Event Index FE	N	Y	N	Y	N	N	N	N
Event Index × Time Interval FE	Y	N	Y	N	Y	Y	Y	Y
Obs	1,003,664	70,890	1,008,421	81,888	1,008,421	1,008,421	85,210	85,210
Adjusted R2	0.060	0.309	0.128	0.322	0.692	0.465	0.565	0.638
Panel E. Pseudo outages								
Outage	0.031	-0.102	0.250	-0.002	0.000	-0.096	-0.023	-0.009
	(1.26)	(-1.40)	(0.92)	(-1.25)	(0.09)	(-0.45)	(-1.02)	(-1.15)
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	N	Y	N	N	N	N	N	N
Event Index FE	N	Y	N	Y	N	N	N	N

Event Index × Time Interval FE	Y	N	Y	N	Y	Y	Y	Y
Obs	1,557,934	89,025	1,561,693	100,606	1,561,693	1,561,693	104,286	104,286
Adjusted R2	0.063	0.306	0.115	0.325	0.722	0.532	0.576	0.730

This table reports supplementary robustness tests for the main 5-minute regressions, including: market-adjusted turnover (*ΔTO*), overall systematic price synchronicity (*Overall SYNCH*), absolute retail order imbalance *(Abs_ROI*), retail order-imbalance volatility (*ROIV*), effective spread, return volatility (*Volatility*), idiosyncratic volatility, and industry-level systematic volatility. As described in Section 4.1, each panel re-estimates the corresponding baseline specification under an alternative identification screen: Panel A excludes firm-specific news-event observations; Panel B removes outage events that overlap with scheduled macroeconomic announcements (CPI, PPI, and FOMC minutes) released before or during the outage window; Panel C excludes outages shorter than 30 minutes; Panel D excludes outages that begin before 10:00 AM; and Panel E implements a placebo by assigning the outage indicator to day t−1. Controls, fixed effects and standard errors follow the baseline specifications.

Table C.2: Robustness Tests (B): Pseudo outages

	Return t, [p+1]	Return t, [p+1, p+2]	Return t, [p+1, p+3]	Return t, [p+1, p+5]	Return [t+1]	Return [t+1, t+3]	Return [t+1, t+5]
Panel A. Using Pseudo outages for Table 8							
Outage x Selling ROI	0.003	0.006	0.011	0.035	0.222	0.313	0.507*
	(0.59)	(0.66)	(0.88)	(1.63)	(1.13)	(1.22)	(1.74)
Outage x Buying ROI	-0.003	-0.006	-0.008	-0.020	-0.086	-0.035	0.097
	(-0.53)	(-0.57)	(-0.50)	(-0.76)	(-0.39)	(-0.17)	(0.33)
Selling ROI	0.006**	0.007*	0.003	-0.005	-0.025	-0.133*	-0.195**
	(2.52)	(1.74)	(0.70)	(-0.73)	(-0.44)	(-1.71)	(-2.25)
Buying ROI	0.009***	0.013***	0.019***	0.022**	0.070	0.058	0.170
	(3.58)	(3.29)	(3.38)	(2.23)	(1.05)	(0.57)	(1.48)
Outage	0.001	0.004	0.008	0.022	-0.181	-0.192	-0.171
	(0.11)	(0.41)	(0.46)	(0.74)	(-0.67)	(-0.53)	(-0.35)
Controls	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y
Event Index FE× Time Interval FE	Y	Y	Y	Y	Y	Y	Y
Obs	1,549,585	1,494,659	1,406,620	1,238,014	1,561,653	1,561,653	1,561,653
Adjusted R2	0.022	0.027	0.031	0.039	0.103	0.204	0.242
Panel B. Using Pseudo outages for Table 9							
Outage x ROI_other	0.007	0.012	0.014	0.033*	0.317	0.213	0.429
	(1.22)	(1.49)	(1.35)	(1.89)	(1.47)	(0.82)	(1.59)
Outage x ROI_liquidity	0.008	0.018	0.021	0.007	0.279	-0.010	-0.229
	(0.48)	(0.80)	(0.72)	(0.12)	(0.85)	(-0.02)	(-0.51)
Outage x ROI_persistence	0.010	0.025*	0.051**	0.056	0.382	0.669	1.006*
	(0.91)	(1.72)	(2.57)	(1.52)	(1.10)	(1.37)	(1.77)
ROI_other	0.013***	0.015***	0.014***	0.014	-0.076	-0.103	-0.077
	(4.97)	(3.66)	(2.70)	(1.65)	(-1.24)	(-1.30)	(-0.95)

ROI_liquidity	-0.083	0.048	-1.337	-1.550	20.998	48.472*	58.527**
	(-0.08)	(0.02)	(-0.43)	(-0.31)	(1.12)	(1.69)	(2.19)
ROI_persistence	0.002	0.005	-0.003	-0.011	-0.011	-0.160	0.134
	(0.34)	(0.51)	(-0.22)	(-0.47)	(-0.09)	(-0.80)	(0.58)
Outage	Y	Y	Y	Y	Y	Y	Y
	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y
Firm FE	584,589	558,693	522,979	454,517	584,580	584,580	584,580
Event Index × Time Interval FE	0.032	0.037	0.042	0.053	0.123	0.214	0.252

This table provides placebo tests for the return-predictability regressions reported in Tables 8 and 9. In both panels, we reassign the outage indicator to the matched pre-event day ($t-1$) and re-estimate the baseline specifications using the same set of controls, fixed effects and standard errors.

Appendix D: Principal Component Analysis

Following Kaniel et al. (2008, KST), we quantify the extent of industry-level and market-level systematic co-movement in 5-minute stock returns (ROI) by extracting common factors from the cross-section of returns (ROI) and benchmarking the resulting principal components against a simulated "no-common-factor" null.

Below we describe the industry-level PCA for 5-minute stock returns; the ROI procedure is identical.

Step A. Construct the industry return matrix

For each industry (g) (Fama–French 12) and each outage event (e), we separately consider (i) the outage window and (ii) the matched control windows. For each sample, we form a time-by-stock matrix:

$$X^{(g,e)} \in R^{TxN}$$

where rows index 5-minute time bins within the event window (stacked across days for controls), columns index stocks (i= 1, …, N) in industry (g), and each element is the 5-minute return. Missing entries are set to zero, and stocks with zero time-series variance over the window are excluded.

Step B. PCA on the correlation matrix (Real)

To ensure that PCA captures co-movement rather than scale differences, we standardize each stock's return series:

$$Z_{i,t} = \frac{X_{i,t} - \overline{X_\iota}}{S_i}$$

where $\overline{X_\iota}$ and S_i are the sample mean and standard deviation of stock (i) in the window. We then run PCA on Z. Let λ_k denote the k-th eigenvalue of the correlation matrix. The percentage of variance explained by principal component k is:

$$RealEVR_k = \frac{\lambda_k}{\sum_{j=1}^{\min(T,N)} \lambda_j}, \qquad k = 1,2$$

Step C. Simulated benchmark (Sim)

Because finite samples can mechanically generate nontrivial leading eigenvalues even under independence, we construct a benchmark as in KST. For each stock (i), we estimate its mean and standard deviation (μ_i, σ_i) from X, then generate an artificial matrix under cross-stock independence:

$$X_{i,t}^{sim} \sim N(\mu_i, \sigma_i^2)$$

independently across i. We repeat this simulation 10 times and compute $SimEVR_k$ for each draw using the same correlation-PCA procedure as in Step B; $SimEVR_k$ is the average across the 10 simulations.

Step D. Systematic component (Real − Sim)

We define the systematic variance share of component (k) as the benchmark-adjusted explained variance:

$$Systematic_k = RealEVR_k - SimEVR_k, \qquad k = 1,2$$

Step E. Random subsampling and aggregation

To mitigate sensitivity to the cross-sectional dimension, we implement KST-style subsampling. For each (g,e) sample, we repeatedly draw K=min(30,N) stocks without replacement, compute $Systematic_k$ for k=1,2, and average across M=100 subsamples:

$$\overline{Systematic_k} = \frac{1}{M}\sum_{m=1}^{M} Systematic_k^m$$

The PCA for ROI procedure is identical.

The market-level procedure is similar, except that the time-by-stock matrix $X^{(e)}$is constructed using all stocks in the sample rather than restricting to industry g.