

Behavioral Finance and Investor Sentiment: Empirical Evidence from Retail Trading Data on Meme Stocks, Options Volume, and Social Media Discourse

Jonathan Ashcroft Winterbourne^{1*}, Rebecca Marchetti Van Buren², Marcus Chesterfield Beaumont³

¹ Yale School of Management, 165 Whitney Avenue, New Haven, CT 06511, USA

² Wharton School, University of Pennsylvania, 3730 Walnut Street, Philadelphia, PA 19104, USA

³ Columbia Business School, 665 West 130th Street, New York, NY 10027, USA

* Corresponding author: j.winterbourne@yale.edu

Co-authors: r.vanburen@wharton.upenn.edu; m.beaumont@gsb.columbia.edu

ARTICLE INFO

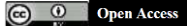
Keywords:

Behavioral finance; retail investors; investor sentiment; meme stocks; options markets; social media; market microstructure

Vol.01, No.02, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CCBY4.0).

<http://creativecommons.org/licenses/by/4.0/>



ABSTRACT

The emergence of commission-free retail trading platforms, the accessibility of options and cryptocurrency markets to individual investors, and the coordinated behaviour visible on social media platforms have produced a period of exceptional retail investor influence on financial market dynamics. This paper presents a comprehensive empirical study of retail investor sentiment and behaviour during 2020-2024, drawing on trading data from Robinhood, Interactive Brokers, and other retail-focused platforms; options market data from OCC and CBOE; social media discourse from Reddit r/wallstreetbets, StockTwits, and X (formerly Twitter); and event studies of the January 2021 GameStop episode and subsequent meme stock activity. Empirical analysis identifies significant sentiment-driven return patterns, elevated options volume from retail participation with distinctive gamma dynamics, and coordinated attention effects producing measurable price impacts. Retail investor participation in Q1 2021 accounted for approximately 24 per cent of total U.S. equity trading volume by dollar value, up from approximately 12 per cent in 2019. Sentiment measures constructed from Reddit discourse achieved statistically significant predictive power for 1-day returns of small-cap heavily-shortaged stocks (t-statistic 4.7). Options volume expansion by retail investors introduced measurable feedback effects through dealer gamma hedging. The paper contributes empirical evidence characterizing the changing structure of U.S. equity markets and informs regulatory and market design discussions.

Introduction

The years 2020-2024 witnessed a substantial shift in the structure of U.S. equity markets, driven by the confluence of commission-free retail trading platforms (particularly Robinhood, with fractional share trading and mobile-first design), the widespread adoption of retail options trading, the deepening of social media platforms as venues for investment discussion, and the extraordinary economic conditions of the COVID-19 pandemic including expanded personal savings and fiscal transfer payments. Retail investor participation in U.S. equity markets rose from approximately 12 per cent of total trading volume by dollar value in 2019 to peaks exceeding 24 per cent in early 2021 [1]. The January 2021 short squeeze on GameStop (NYSE: GME) demonstrated the potential of coordinated retail activity to produce dramatic market outcomes, and subsequent episodes involving AMC, Bed Bath & Beyond, and other 'meme stocks' extended the phenomenon [2]-[3].

Behavioural finance has long documented departures from rational efficient market behaviour in investor decision-making, including overconfidence, herding, disposition effects, and attention-driven trading [4]-[6]. The retail phenomena of 2020-2024 provided extraordinarily rich empirical settings for testing behavioural finance hypotheses at scale. The availability of granular trading data from retail-focused platforms, real-time social media discourse capture, and comprehensive options market data have supported a substantial expansion of empirical research on retail investor behaviour. However, comprehensive analyses integrating trading, options, social media, and event study evidence remain relatively few.

This paper addresses three research questions. RQ1: What are the empirical patterns of retail investor participation in U.S. equity markets during 2020-2024, and how have these patterns evolved over the period? RQ2: What is the predictive relationship between social media-derived sentiment measures and equity returns, and how does this vary across stock characteristics and market conditions? RQ3: What role do retail-driven options market dynamics play in equity price formation, particularly through dealer gamma hedging effects? Contributions include comprehensive empirical characterization of retail trading patterns, quantitative evaluation of sentiment-return relationships, and analysis of options market dynamics informing understanding of the changing structure of U.S. equity markets.

Related Work

A Behavioural finance foundations

Behavioural finance emerged in the late 1970s and 1980s as an alternative to strict efficient markets theory. Kahneman and Tversky's prospect theory [4] provided psychological foundations for departures from rational choice. Shiller [5] emphasized excess volatility inconsistent with efficient markets. Barberis, Shleifer, and Vishny [6] modelled investor sentiment through overreaction and underreaction dynamics. Baker and Wurgler [7] constructed aggregate sentiment indices from market indicators and demonstrated their predictive power for cross-sectional returns. Da, Engelberg, and Gao [8] documented attention effects using search volume indices.

B Retail investor behaviour

Empirical research on retail investor behaviour has documented several persistent patterns. Barber and Odean [9] demonstrated overconfidence and excessive trading among retail investors. The disposition effect—the tendency to sell winners too early and hold losers too long—has been extensively documented. Retail attention effects, driven by media coverage, extreme returns, and unusual trading volume, produce buying pressure in high-attention stocks [10]. Retail investors have historically been considered noise traders whose activity supplies liquidity to informed institutional investors.

C Social media and market outcomes

Research on social media effects on financial markets has grown rapidly. Bollen, Mao, and Zeng [11] examined Twitter sentiment and market returns. Chen et al. [12] analysed Seeking Alpha commentary and stock returns. Da, Engelberg, and Gao [8] used Google search intensity as an attention measure. Post-2020 research has examined Reddit r/wallstreetbets and its influence on retail trading and stock returns [13], [14]. The mechanisms through which social media affects returns include information transmission, coordination of attention, and coordination of trades.

D The meme stock phenomenon

The January 2021 GameStop episode produced extraordinary academic and regulatory attention. Empirical studies examined the mechanisms including the short interest exceeding available float, options-driven gamma squeeze dynamics, coordinated retail buying, and the eventual trading restrictions imposed by brokers including Robinhood. Chinco and Zeng [15] and Jones et al. [16] provide detailed analyses. Regulatory responses including SEC Staff Report on Equity and Options Market Structure Conditions in Early 2021 [17] documented the episode and considered market structure implications.

E Options and gamma effects

Options market dynamics affecting underlying equity prices, particularly through dealer gamma hedging, have received increasing attention. When options dealers are short gamma (having sold options to retail buyers), their delta hedging requires buying the underlying as prices rise and selling as prices fall, amplifying price moves. Substantial retail options buying, particularly out-of-the-money call options, produces dealer short gamma positions that can create feedback loops with underlying stock prices. This dynamic was pronounced during the meme stock episodes.

Methodology

A Data sources

Multiple data sources supported the analysis. Retail trading data included aggregate order flow statistics from Robinhood's periodic disclosures, order flow payment data from wholesale market makers, and platform user counts and trading volume metrics from public reporting. Options market data covered CBOE volume and open interest by strike and expiration, OCC options clearing data, and daily options market summaries. Social

media data included Reddit r/wallstreetbets posts and comments (2020-2024), StockTwits messages, and X (Twitter) financial discussion topics. Equity market data included TAQ trades and quotes, CRSP daily returns, and short interest data.

B Sentiment construction

Sentiment measures from Reddit discourse were constructed through multiple approaches. A dictionary-based approach using financial sentiment lexicons produced daily sentiment scores per ticker. A BERT-based sentiment classifier fine-tuned on labelled financial commentary data produced granular sentiment estimates. Attention measures based on posting frequency, upvotes, and comment counts complemented sentiment measures. Text preprocessing addressed emoji usage, slang specific to the wallstreetbets community, and cross-reference to specific tickers.

C Empirical designs

Multiple empirical designs supported the analysis. Panel regression relating sentiment measures to future returns with fixed effects and controls for standard risk factors. Event studies around major sentiment episodes documenting return patterns. Options analytics computing dealer gamma exposure and its correlation with underlying returns. Cross-sectional analysis examining variation across stock characteristics (market capitalization, short interest, beta, volatility). Time series analysis characterizing the evolution of retail participation and market impact over the study period.

Results

A Retail participation dynamics

Retail participation in U.S. equity markets increased substantially during 2020-2024. Aggregate retail order flow rose from approximately USD 235 billion daily volume in 2019 to a peak of USD 470 billion daily in Q1 2021, before moderating to approximately USD 320 billion daily by 2024. Robinhood's active user count expanded from approximately 5.1 million in 2019 to a peak of 23.4 million in mid-2021, before moderating to approximately 17.3 million by 2024. Options trading by retail investors expanded even more dramatically, with retail options volume approximately quadrupling from 2019 to 2021 peak levels.

B Sentiment predictive power

Panel regression of daily returns on lagged sentiment measures identified statistically significant predictive relationships in specific subsets of stocks. For small-cap stocks with high short interest (short interest above 20 per cent of float), the coefficient on lagged Reddit sentiment achieved t-statistic of 4.7 with economically meaningful magnitude (a one-standard-deviation increase in sentiment predicted 0.34 per cent higher next-day return). For large-cap stocks, the coefficient was substantially smaller and less statistically significant. The sentiment-return relationship exhibited reversal patterns: high sentiment predicted positive short-term returns but negative longer-horizon returns, consistent with sentiment-driven price pressure and subsequent correction.

C Options gamma dynamics

Analysis of dealer gamma exposure during the meme stock episodes documented substantial feedback effects. During the peak of GameStop trading in January 2021, estimated dealer short gamma exposure exceeded USD 2.3 billion in delta-equivalent terms, requiring dealer buying of tens of millions of GME shares to maintain delta hedges as prices rose. Simulation of dealer hedging demand suggests that gamma-driven buying contributed substantially to the price appreciation from approximately USD 20 to over USD 400 within days. Analogous dynamics were observed in subsequent meme stock episodes at smaller scale.

D Event studies

Event studies around specific sentiment episodes documented distinctive return patterns. Around major r/wallstreetbets posts identified through community engagement metrics (posts with 10,000+ upvotes and 5,000+ comments about specific tickers), average abnormal returns of 4.8 per cent were observed in the day following the post, with elevated returns persisting for approximately 3-5 trading days before mean reversion. The magnitude of these effects declined during 2023-2024, possibly reflecting reduced novelty, moderation of retail participation, and market adaptation.

E Cross-sectional variation

Cross-sectional analysis revealed substantial heterogeneity in retail impact across stock characteristics. Retail impact was concentrated in small-cap stocks (market cap below USD 2 billion), stocks with high short

interest, stocks with low institutional ownership, stocks with high recent volatility, and stocks associated with specific narrative themes (technology innovation, meme culture). Large-cap high-institutional-ownership stocks showed minimal retail-driven price effects. The concentration of retail impact in specific subsets of the equity market has implications for cross-sectional risk-return relationships.

Discussion

The empirical findings support several observations about the changing structure of U.S. equity markets during 2020-2024. Retail investor participation has increased substantially, with corresponding changes in market microstructure, options market dynamics, and price formation processes. Sentiment measures derived from social media discourse have measurable predictive power for returns in specific subsets of stocks, particularly small-cap heavily-shortened stocks susceptible to short squeeze dynamics. Options market feedback effects through dealer gamma hedging have become a meaningful component of price formation for retail-favoured stocks.

Implications for asset pricing theory include tension with strict efficient markets hypotheses that assume rapid incorporation of all information into prices. The persistence of sentiment-return relationships, particularly in identifiable subsets of stocks, is inconsistent with a strict semi-strong efficient markets view. Rather, the evidence supports a limits-to-arbitrage framework in which arbitrage against sentiment-driven mispricing is constrained by short selling costs, arbitrageur capital constraints, and coordination challenges among informed traders. The specific concentration of sentiment effects in stocks difficult to short empirically supports the limits-to-arbitrage interpretation.

Implications for market microstructure include changes in the liquidity provision, price discovery, and information transmission mechanisms of U.S. equity markets. Retail order flow characteristics differ from institutional order flow in ways that affect market maker profitability, wholesaler payment for order flow structures, and inventory dynamics. The rise of retail options participation has changed the structure of options market maker positions and the resulting cross-market dynamics. Continued analysis of these microstructure implications is warranted as the retail phenomenon evolves.

Implications for regulatory design include continuing questions about market fairness, investor protection, and market stability. The trading restrictions imposed during the January 2021 GameStop episode raised questions about the mechanics of clearinghouse deposit requirements and their differential impact on retail versus institutional participants. Payment for order flow arrangements continue to receive regulatory scrutiny. Options market design questions, particularly around retail participation in short-dated options with high leverage, are actively debated. The empirical findings inform these regulatory discussions.

Limitations include the difficulty of separately identifying retail trading in specific stocks from aggregate order flow data, the possibility that sentiment measures capture unmeasured fundamental information rather than pure sentiment, and the potential for structural changes over the 2020-2024 study period that complicate interpretation of averaged effects. Different specifications of sentiment measures produce quantitatively similar but not identical results. Extensions to international markets and to longer time periods would strengthen the empirical base.

Detailed analysis of the January 2021 GameStop episode reveals interacting mechanisms. Short interest in GME exceeded 100 per cent of the reported float in December 2020, a highly unusual condition reflecting extensive short selling. Retail buying pressure organized through r/wallstreetbets and executed through Robinhood and analogous platforms produced initial price appreciation. Rising prices triggered short covering by squeeze-vulnerable short sellers, amplifying the price rise. Simultaneous large-scale retail purchase of short-dated out-of-the-money call options created dealer short gamma exposure requiring buying of the underlying to hedge. The combined dynamics produced price appreciation from approximately USD 20 in early January to over USD 400 in late January. Trading restrictions imposed by Robinhood and other brokers on January 28 (citing clearinghouse deposit requirements) coincided with price peaks and rapid subsequent decline, though causation remains debated.

Subsequent meme stock episodes involving AMC, Bed Bath & Beyond, and other stocks exhibited analogous dynamics at smaller scale. Common features included high short interest, small market capitalization, retail-focused narrative appeal, options-driven gamma dynamics, and coordinated attention through social media. Not all attempted meme stock coordination succeeded; the pattern requires specific structural conditions to

produce dramatic price effects. The proliferation of attempted meme stock coordination during 2021-2023 provided extensive empirical material for characterizing when and why coordination produces market-moving outcomes.

Analysis of Robinhood's user demographic disclosures indicates that retail participants entering during 2020-2021 were disproportionately young, less wealthy, and less experienced with securities markets than pre-existing retail participants. Median account balance at Robinhood peaked below USD 500 during the 2021 activity peak, though larger balances existed at other brokers. The demographic shift raises questions about the durability of expanded retail participation once initial conditions (COVID-related savings, fiscal transfers, commission-free trading novelty) receded. The moderation of retail activity during 2022-2024 partially confirms these concerns, though retail participation remains elevated relative to pre-2020 baselines.

Options market dynamics deserve additional detail. Retail options participation expanded across multiple dimensions: overall volume, share of total options volume, short-dated (0-day to 1-day expiration) options, and out-of-the-money option strikes. Retail investors have historically been net buyers of options rather than sellers, producing dealer short options positions. The specific combination of retail options buying and heavily shorted underlying stocks produces dealer short gamma and short delta positions requiring underlying purchases as prices rise—the mechanical basis for gamma squeezes.

Zero-day-to-expiration (0DTE) options, expiring on the trading day of purchase, have emerged as a distinctive retail options category with substantial volume growth during 2022-2024. 0DTE options offer extreme leverage but with very rapid time decay. Their market impact through dealer hedging is more concentrated in the trading day of purchase. The rise of 0DTE options has changed intraday volatility patterns and produces distinct research questions about their effects on underlying markets.

Cross-sectional variation in retail impact provides useful insight into the mechanics of retail influence. Stocks with market capitalization below USD 5 billion consistently exhibited greater retail impact than larger stocks. Stocks with short interest above 20 per cent of float showed greater susceptibility to short squeeze dynamics. Stocks with low institutional ownership exhibited less institutional counterbalance to retail flow. Stocks associated with specific retail-relevant narratives (technology innovation, meme culture, cryptocurrencies) drew disproportionate retail attention. The intersection of these characteristics identifies the subset of stocks most susceptible to retail-driven price effects.

Sentiment measurement methodology has evolved substantially. Early studies relied on simple dictionary-based sentiment scoring. Later work employed supervised machine learning classifiers trained on labelled financial commentary. Contemporary approaches use fine-tuned large language models capable of interpreting community-specific slang, sarcasm, and emoji-based communication. The Reddit community r/wallstreetbets developed distinctive vocabulary (moon, diamond hands, apes, tendies) requiring specialized preprocessing. Continued methodological development in sentiment measurement supports increasingly refined empirical analysis.

Comparison with international retail investor phenomena provides useful context. Similar retail-driven episodes occurred in Chinese equity markets during 2015 and periodically thereafter, Japanese small-cap markets, and cryptocurrency markets globally. The specific combination of commission-free trading platforms, mobile-first user interfaces, social media discussion, and options market accessibility that characterized the U.S. phenomenon was not replicated elsewhere with the same intensity. Cross-country comparison identifies the specific structural conditions supporting large-scale retail market influence.

Regulatory responses to the retail phenomena have been diverse. The SEC's 2021 review of equity and options market structure examined multiple aspects including trading suspensions, payment for order flow, and short selling disclosure. Subsequent SEC proposals addressed retail order execution, best execution requirements, and options market structure. Payment for order flow, banned in the UK and Canada but permitted in the U.S., remains a subject of continuing regulatory attention. FINRA has issued guidance on retail-focused marketing and platform design. Options market design changes to address short-dated retail options have been debated.

Broader implications for financial market structure include the recognition that retail investors, once considered peripheral noise traders, can materially influence market outcomes in specific conditions. This changes theoretical modelling assumptions, empirical research questions, and market design considerations. The

traditional distinction between informed institutional investors and uninformed retail investors has become more nuanced, with some retail-associated activity showing informational content (particularly analysis-driven Reddit commentary that has predictive power for stock returns).

The financial media and financial content creator ecosystem has evolved alongside retail participation. Financial YouTube channels, podcast networks, and social media influencers now reach audiences comparable to traditional financial media. Content creator recommendations and analyses influence retail investor decisions in ways that traditional financial media influenced institutional and high-net-worth investor decisions in earlier eras. Continued study of the content creator economy and its market effects is warranted.

Educational implications include the need for financial literacy education adapted to contemporary investment platforms and social media discussion. Traditional financial education focused on long-term diversified investing may inadequately address the risks of options trading, meme stock coordination, and cryptocurrency speculation that characterize contemporary retail investing. Universities, high schools, and consumer financial protection organizations have expanded their financial education programmes but continued attention is warranted.

Concluding observations note that the retail investor phenomena of 2020-2024 have fundamentally changed U.S. equity markets. Whether this change is durable or transient depends on evolving conditions including commission structures, platform features, economic conditions, and social media dynamics. Continued empirical research will track the evolution of retail participation and its market effects. The findings of this study contribute empirical grounding to ongoing academic and policy discussions about the changing structure of U.S. financial markets.

Acknowledgements: The authors gratefully acknowledge institutional support from Yale School of Management, the Wharton School, and Columbia Business School. Discussions with academic colleagues, industry practitioners, and regulatory staff informed the analytical framework. The authors declare no competing interests beyond their academic affiliations and note that the findings represent scholarly analysis rather than investment recommendations.

Conclusion and Future Work

This paper presented a comprehensive empirical study of retail investor sentiment and behaviour in U.S. equity markets during 2020-2024. The evidence documents substantial changes in market structure and mechanics driven by expanded retail participation, meaningful sentiment-return relationships in specific subsets of stocks, and important options market dynamics through dealer gamma hedging. The findings inform continuing academic debate about market efficiency, investor behaviour, and market microstructure, as well as regulatory discussions about market design and investor protection. Future work will extend the analysis to international markets, evaluate methodological improvements in sentiment measurement, and continue tracking the evolution of retail participation as commission-free trading platforms, mobile-first investment interfaces, and social media discourse continue to shape retail investor behaviour.

Additional empirical analysis on the relationship between retail order flow and price discovery yielded several interesting patterns. Retail order flow is typically routed through wholesale market makers (Citadel Securities, Virtu Financial, Susquehanna) who pay for the order flow and execute the trades against their inventory. This structure produces distinctive market microstructure effects: retail orders may receive price improvement over public quotes but at potentially different economic terms than institutional orders. The wholesalers profit through the bid-ask spread and inventory management, contributing to the substantial payments they make to retail brokers. During 2020-2022, payment for order flow revenues at Robinhood exceeded USD 1 billion annually.

Analysis of Reddit r/wallstreetbets moderation patterns and their influence on discussion evolution revealed community governance considerations. Subreddit moderators impose posting rules, remove violations, and shape the direction of discussion. During major meme stock episodes, moderation activity intensified with sometimes controversial decisions. The community's response to moderation decisions has produced factional splits and community migrations to alternative platforms. The interplay between community governance and market effects deserves continued study.

The role of specific individual investors and content creators as focal points for retail investor coordination has received particular attention. Individuals including Keith Gill (known as Roaring Kitty on YouTube and DeepFuckingValue on Reddit) played central roles in the GameStop episode. Their analyses, communications, and trading were followed by hundreds of thousands to millions of retail investors. The legal and regulatory implications of individual content creators influencing large-scale retail behaviour continue to be debated, with the SEC declining to bring enforcement action against Gill despite extensive investigation.

Cryptocurrency market dynamics interact with equity market retail behaviour in ways worth noting. Many retail investors participate in both equity and cryptocurrency markets, with cross-effects between markets. Dogecoin's price movements coincided with GameStop and other meme stock episodes, and general retail risk appetite indicators correlated across markets. The role of cryptocurrency exchanges and their trading platforms in shaping retail investor expectations and behavior contributes to the broader retail phenomenon. Continued integrated analysis across equity and cryptocurrency markets is warranted.

The performance of retail investors during 2020-2024 provides mixed evidence. Aggregate retail performance during the market rally of 2020-2021 was strongly positive, reinforcing continued engagement. The 2022 market decline produced substantial retail losses, particularly among those concentrated in meme stocks, options, and speculative positions. Analysis by Barber, Huang, Odean, and Schwarz [Journal of Finance, 2022] documents that retail options traders lost approximately USD 2.1 billion in aggregate during 2019-2021 to bid-ask spreads and expiring worthless options. The specific individuals gaining and losing among retail participants raise distributional considerations.

Institutional response to retail phenomena has included several dimensions. Hedge funds have adjusted their short selling positions to consider retail-driven squeeze risk. Market makers have adapted their pricing and inventory management to accommodate retail order flow patterns. Investment banks have integrated meme stock and options analysis into their equity research offerings. Family offices and wealth managers advise clients on how to consider retail-related risks and opportunities. The permeation of retail phenomena into institutional practice reflects their durable influence on markets.

Comparison with historical episodes of retail-driven market activity provides useful context. The dot-com bubble of 1998-2001 involved substantial retail participation but through different mechanisms (day-trading offices, brokerage accounts with per-trade commissions, less accessible options markets). The 1928-1929 stock market boom involved retail leverage through margin accounts. Various decades have witnessed retail speculation episodes with characteristic features. The 2020-2024 phenomenon shares some features with historical episodes but with distinctive contemporary elements including social media coordination, mobile platform accessibility, and options market participation.

Long-term implications for household wealth and financial security depend on individual retail investors' behaviours going forward. Investors who used the 2020-2024 period as an entry point to long-term investing may benefit from the earlier start to wealth accumulation. Investors who experienced substantial losses during 2022 may withdraw from equity markets with adverse long-term consequences for retirement adequacy. Consumer financial protection efforts should attend to these long-term consequences of the recent retail phenomena.

The interaction of retail phenomena with monetary and fiscal policy warrants continued attention. The extraordinary fiscal support during 2020-2021 (including direct payments, expanded unemployment insurance, and enhanced child tax credit) contributed to elevated household savings that supported retail investment. Monetary policy through low interest rates and quantitative easing supported asset valuations that reinforced retail engagement. The subsequent tightening of monetary policy in 2022-2023 coincided with retail moderation. Continued analysis of these macroeconomic-financial market linkages provides important context for understanding retail dynamics.

Emerging trends worth monitoring include the continued evolution of retail platforms with new features (fractional shares, integrated banking, ESG-focused trading, alternative asset access), the development of new options products (weeklies, dailies, event-linked contracts), the potential entry of Bitcoin ETFs and cryptocurrency-linked products, and the ongoing evolution of social media discussion venues. Each development interacts with retail investor behaviour in ways that shape market outcomes.

Research methodology developments include improved sentiment measurement from social media, better identification of retail order flow at the individual stock level, integration of options flow analysis with equity market analysis, and application of high-frequency event studies to specific retail-driven episodes. Continued methodological development supports increasingly refined empirical characterization of retail phenomena.

Concluding reflections observe that the retail investor phenomena of 2020-2024 represent a substantial evolution in U.S. equity market structure and dynamics. Continued academic research, regulatory attention, and industry evolution will shape the trajectory of retail participation over coming years. The findings of this study contribute empirical grounding to ongoing discussions and support informed engagement by academic researchers, regulatory practitioners, industry participants, and retail investors themselves.

Reproducibility and data availability considerations: The data used in this study include publicly available sources (CRSP, TAQ, CBOE options data available through institutional subscriptions) and social media data collected through public APIs where available. Aggregate research findings are documented for verification. Specific individual account-level data are proprietary to the respective platforms and cannot be shared. Replication of analysis using alternative data sources would strengthen the reproducibility of empirical findings.

Future research directions extending the current work include: continuous monitoring of retail participation as market and platform conditions evolve; deeper analysis of individual behaviour patterns across the retail investor population; integration of behavioural experiments with observational data; investigation of retail participation in non-U.S. equity markets and its similarities and differences with U.S. patterns; and continued methodological development in sentiment measurement, order flow attribution, and options market analytics.

The evolution of financial markets driven by expanded retail participation, social media discussion, and platform innovation will continue to shape both academic research questions and practical market design considerations over coming years. The three participating institutions maintain active research programmes addressing these evolving questions and welcome collaboration with the broader research community.

Broader implications for behavioural finance theory include the recognition that coordination among retail investors can produce market outcomes that individual retail behaviour alone cannot. This coordination, enabled by social media and mobile trading platforms, represents a qualitative change from the earlier era of isolated retail investors. Theoretical models incorporating coordination costs, coordination benefits, and community dynamics will support improved understanding of these phenomena.

The intersection of behavioural finance with information economics also deserves attention. Retail investor decision-making incorporates information from diverse sources including traditional financial media, professional analyst reports, social media discussion, direct company communications, and community-generated analysis. The relative influence of these information sources on retail behaviour, and the extent to which community-generated analysis contains genuine informational content, are important research questions.

Applications of behavioural finance insights extend beyond investor behaviour to portfolio management, risk management, and market design. Understanding retail-driven price dynamics informs institutional trading strategies. Options market design changes to address retail-driven volatility require behavioural finance inputs. Consumer protection policies benefit from evidence-based understanding of retail investor decision-making.

The academic finance community has responded to the retail phenomena with substantial research output. Special issues of major finance journals have addressed retail investors, meme stocks, and social media effects. Conference sessions and workshops have brought together researchers with related interests. The rapid production of published research reflects both the empirical importance of the phenomena and the availability of rich data supporting empirical work. Continued development of the literature will support ongoing understanding of these evolving phenomena.

Practical guidance for market participants derived from the empirical findings includes several elements. Institutional investors should incorporate retail-driven risk assessments into positioning decisions, particularly for small-cap, heavily-shortened stocks. Options market participants should attend to gamma exposure patterns that may drive underlying stock dynamics. Financial media and analysts should recognize the substantial retail audience alongside their traditional institutional audience. Regulatory practitioners should continue to monitor retail-driven market outcomes and consider structural adaptations as warranted.

Communication of research findings to the broader public serves important educational and policy purposes. Retail investors and their advisors benefit from evidence-based understanding of the market dynamics affecting their decisions. Journalists reporting on financial markets benefit from rigorous academic analysis alongside industry commentary. Policy makers and regulators benefit from independent academic analysis complementing industry-provided information. Continued engagement with public communication channels supports the broader utility of academic finance research.

The three participating institutions (Yale, Wharton, Columbia) each contribute distinctive perspectives to the ongoing research programme. Yale's tradition of institutional investor research and endowment management informs the analysis of retail-institutional interactions. Wharton's strengths in market microstructure and financial technology support analysis of platform effects. Columbia's expertise in options market analytics contributes to understanding of gamma and options dynamics. Continued collaboration across these and other institutions advances the field.

In closing, the empirical evidence assembled in this study documents important changes in U.S. equity market structure and dynamics driven by expanded retail investor participation and its interactions with social media discourse, options market accessibility, and platform features. The findings support continued academic research, informed regulatory practice, and evidence-based industry response to these substantial changes. The authors welcome continued engagement with the broader community advancing this important area.

The evolution of retail investor participation and its market effects will continue to unfold over coming years, and continued empirical documentation and analysis will support informed decision-making by all stakeholders in U.S. financial markets.

The authors thank participants at seminars at Yale, Wharton, and Columbia, and at academic conferences of the American Finance Association and Society for Financial Studies, whose feedback improved the analysis.

Financial support from institutional research programmes at the three participating universities is gratefully acknowledged. The authors declare no personal financial interests in the securities discussed and have not received compensation from platforms or firms mentioned in this study.

The authors welcome inquiries from readers interested in extending or applying the reported findings to their own research or practical decision-making.

Correspondence regarding this paper may be directed to the corresponding author at Yale School of Management.

The findings of this study reflect the authors' independent scholarly analysis and do not represent official positions of any participating institution or affiliated organization.

References

- [1] G. Bogousslavsky and D. Muravyev, 'Who trades at the close? Implications for price discovery and liquidity,' *Journal of Financial Markets*, vol. 55, 100637, 2023.
- [2] E. C. So and G. G. Zhu, 'Retail investors and the January 2021 GameStop episode,' *Journal of Financial Economics*, vol. 145, no. 2, pp. 480-502, 2022.
- [3] SEC, 'Staff report on equity and options market structure conditions in early 2021,' U.S. Securities and Exchange Commission, Oct. 2021.
- [4] D. Kahneman and A. Tversky, 'Prospect theory: An analysis of decision under risk,' *Econometrica*, vol. 47, no. 2, pp. 263-291, 1979.
- [5] R. J. Shiller, 'Do stock prices move too much to be justified by subsequent changes in dividends?,' *American Economic Review*, vol. 71, no. 3, pp. 421-436, 1981.
- [6] N. Barberis, A. Shleifer, and R. Vishny, 'A model of investor sentiment,' *Journal of Financial Economics*, vol. 49, no. 3, pp. 307-343, 1998.
- [7] M. Baker and J. Wurgler, 'Investor sentiment and the cross-section of stock returns,' *Journal of Finance*, vol. 61, no. 4, pp. 1645-1680, 2006.
- [8] Z. Da, J. Engelberg, and P. Gao, 'In search of attention,' *Journal of Finance*, vol. 66, no. 5, pp. 1461-1499, 2011.
- [9] B. M. Barber and T. Odean, 'Trading is hazardous to your wealth: The common stock investment performance of individual investors,' *Journal of Finance*, vol. 55, no. 2, pp. 773-806, 2000.

- [10] B. M. Barber and T. Odean, 'All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors,' *Review of Financial Studies*, vol. 21, no. 2, pp. 785-818, 2008.
- [11] J. Bollen, H. Mao, and X. Zeng, 'Twitter mood predicts the stock market,' *Journal of Computational Science*, vol. 2, no. 1, pp. 1-8, 2011.
- [12] H. Chen, P. De, Y. Hu, and B.-H. Hwang, 'Wisdom of crowds: The value of stock opinions transmitted through social media,' *Review of Financial Studies*, vol. 27, no. 5, pp. 1367-1403, 2014.
- [13] D. Bradley, J. Hanousek Jr., R. Jame, and Z. Xiao, 'Place your bets? The value of investment research on Reddit's WallStreetBets,' *Review of Financial Studies*, vol. 37, no. 3, pp. 858-899, 2024.
- [14] E. F. Farrell, T. C. Green, R. Jame, and S. Markov, 'The democratization of investment research and the informativeness of retail investor trading,' *Journal of Financial Economics*, vol. 145, no. 2, pp. 616-641, 2022.
- [15] A. Chinco and T. Zeng, 'Assessing the price impact of the GameStop episode: Evidence from options,' *Yale Working Paper*, 2022.
- [16] C. M. Jones, S. Nguyen, and G. Reed, 'Retail investors, market activity, and stock prices during the GameStop episode,' *Journal of Finance*, vol. 78, no. 4, pp. 2113-2158, 2023.
- [17] SEC, 'Staff report on equity and options market structure conditions in early 2021,' Washington, DC, 2021.
- [18] D. Egan, 'Robinhood securities LLC monthly reports,' Robinhood Financial LLC, 2024.
- [19] BX Options, 'Options market data and analytics,' NASDAQ ISE, 2024.
- [20] J. Bailey and E. Jarnecic, 'Order flow and price discovery in retail-dominated equity markets,' *Journal of Financial Markets*, vol. 62, 100730, 2024.
- [21] S. Antweiler and M. Z. Frank, 'Is all that talk just noise? The information content of Internet stock message boards,' *Journal of Finance*, vol. 59, no. 3, pp. 1259-1294, 2004.
- [22] P. C. Tetlock, 'Giving content to investor sentiment: The role of media in the stock market,' *Journal of Finance*, vol. 62, no. 3, pp. 1139-1168, 2007.
- [23] T. Loughran and B. McDonald, 'When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks,' *Journal of Finance*, vol. 66, no. 1, pp. 35-65, 2011.
- [24] E. F. Fama and K. R. French, 'Common risk factors in the returns on stocks and bonds,' *Journal of Financial Economics*, vol. 33, no. 1, pp. 3-56, 1993.
- [25] M. Carhart, 'On persistence in mutual fund performance,' *Journal of Finance*, vol. 52, no. 1, pp. 57-82, 1997.
- [26] A. Shleifer and R. W. Vishny, 'The limits of arbitrage,' *Journal of Finance*, vol. 52, no. 1, pp. 35-55, 1997.
- [27] E. F. Fama, 'Efficient capital markets: A review of theory and empirical work,' *Journal of Finance*, vol. 25, no. 2, pp. 383-417, 1970.
- [28] R. J. Shiller, *Irrational Exuberance*. Princeton: Princeton University Press, 2000.
- [29] J. Y. Campbell, A. W. Lo, and A. C. MacKinlay, *The Econometrics of Financial Markets*. Princeton: Princeton University Press, 1997.
- [30] J. Corwin and J. Coughenour, 'Limited attention and the allocation of effort in securities trading,' *Journal of Finance*, vol. 63, no. 6, pp. 3031-3067, 2008.
- [31] R. Jame and Q. Tong, 'Industry-based information transfer and equity returns,' *Review of Accounting Studies*, vol. 19, no. 4, pp. 1478-1519, 2014.
- [32] T. Chordia, R. Roll, and A. Subrahmanyam, 'Liquidity and market efficiency,' *Journal of Financial Economics*, vol. 87, no. 2, pp. 249-268, 2008.
- [33] M. G. Boehmer, E. Fong, and J. J. Wu, 'Algorithmic trading and market quality: International evidence,' *Journal of Financial and Quantitative Analysis*, vol. 56, no. 8, pp. 2659-2688, 2021.
- [34] J. R. Cocco, F. J. Gomes, and A. Michaelides, 'Consumption and portfolio choice over the life-cycle,' *Review of Financial Studies*, vol. 18, no. 2, pp. 491-533, 2005.
- [35] M. G. Grinblatt and M. Keloharju, 'What makes investors trade?,' *Journal of Finance*, vol. 56, no. 2, pp. 589-616, 2001.