# Can AI Agents Learn to Misbehave?

# Strategic Exploitation and Financial Stability

Kumushoy Abduraimova* Valerie Laturnus†

August 13, 2026

*Preliminary draft — work in progress. Please do not circulate.*

**ABSTRACT**

This paper investigates the implications of AI trading agents for financial stability. We ask whether such agents can learn to harm or exploit other market participants without being explicitly instructed to do so, and whether such behavior destabilizes markets. We combine regression analysis with trading experiments using large language model and reinforcement learning agents in market environments based on classic studies of asset bubbles and crashes. Across specifications, we find no systematic evidence that harmful, exploitative, or strategic behavior emerges endogenously, nor that market feedback vulnerability or AI trading behavior we study leads to economically meaningful deterioration in financial stability.

**JEL classification**: C63, C92, D83, G01, G12, G14

**Keywords:** AI agents, algorithmic trading, financial stability, reinforcement learning, large language models, market manipulation

*Durham University Business School. ✉ kumushoy.abduraimova@durham.ac.uk
†Durham University Business School. ✉ valerie.laturnus@durham.ac.uk

# 1 Introduction

Can AI agents learn to profit by inducing other market participants to make systematically disadvantageous trading decisions? As autonomous AI trading systems increasingly enter financial markets, this question becomes an important regulatory concern. AI agents are typically instructed to maximize an objective such as trading profits, not to manipulate prices or harm other traders. Yet harmful behavior may emerge instrumentally when exploiting predictable responses by other market participants becomes a profitable way to achieve that objective.

Recent evidence from AI research shows that such behavior need not be explicitly programmed. MacDiarmid et al. (2025) show that when large language models (LLMs) learn to complete tasks by finding unintended shortcuts rather than solving them properly, they may later exhibit other harmful behaviors, such as deception or sabotage, despite never being directly trained to do so.

One setting in which this idea can apply in financial markets is the strategic exploitation of positive feedback traders, i.e., traders who buy after prices rise and sell after prices fall. An AI agent that observes this pattern might learn to exploit it by trading in a way that reinforces the initial price move, prompts further buying or selling, and then unwinds its position at more favorable prices. This type of strategic trading differs from conventional algorithmic collusion because the agent is not coordinating with another strategic trader. Instead, it individually exploits a predictable response by changing the market environment in a way that induces another participant to trade to its advantage.

Dou et al. (2025b) show that such dynamics can emerge in theory. In their environment, reinforcement learning agents with an explicit multi-step planning capability learn to trade aggressively, induce positive feedback investors to amplify the resulting price distortion, and subsequently unwind their positions. The agents are not instructed to manipulate feedback traders and are not programmed with knowledge of their existence. However, the strategy emerges through repeated learning. With strong positive feedback traders presence, planning agents create and exploit temporary price distortions. When feedback trading is limited, strategic bubble creation disappears, and the resulting coordination can instead be mildly

stabilizing.

This study motivates a central distinction in our paper. A market may be susceptible to strategic exploitation without actually being exploited. Strong feedback between initial price movements and subsequent order flow creates an exploitable opportunity, but no instability arises unless agents are able and willing to exploit it. Market vulnerability therefore depends on the trading environment, while its realization depends on the capabilities and incentives of market participants.

Testing this distinction creates an identification problem. High-frequency market data reveal whether earlier price movements induce subsequent trading and whether stronger feedback is followed by greater instability. They do not reveal whether an order originates from an AI planning agent, what objective that agent is pursuing, or whether it is deliberately trying to induce another trader's response. By contrast, controlled experiments allow us to observe agents' information, objectives, actions, and profits directly, but cannot establish how common the relevant vulnerability is in real markets. We therefore combine the two approaches. The market analysis measures the extent of potentially exploitable feedback conditions, while the experiments reveal how autonomous AI agents behave when their presence and objectives are known.

We first estimate an empirical measure of susceptibility to feedback trading, motivated by Dou et al. (2025b). Around precisely timed public information events, we separate the initial price response from subsequent order flow and estimate the sensitivity of the latter to the former, denoted by $\mu$. A larger $\mu$ means greater feedback traders presence, i.e. where an initial price increase is followed by stronger buying pressure, or an initial decline by stronger selling pressure. Importantly, $\mu$ measures the strength of a feedback trading mechanism that a sufficiently capable strategic agent could exploit.

Our preliminary empirical analysis uses trades and quotes data for S&P 500, crude oil, gold, and bitcoin futures around FOMC announcements between 2019 and 2026. We estimate feedback over multiple return and order flow horizons and measure financial instability as the relative realized volatility after the complete feedback estimation interval to its pre-announcement level. Ongoing work extends this analysis to several hundred securities, additional asset classes, and other information events.

Market evidence yields an informative first result. Greater susceptibility to feedback trading is not, by itself, associated with greater subsequent financial instability in the current sample. In our baseline analysis, higher $\hat{\mu}$ is associated with lower subsequent volatility amplification after including fixed effects for the contract, return horizon, and order flow window. This result, however, does not imply that feedback trading is stabilizing. Its effect depends on whether the feedback channel is strategically exploited or not.

We therefore turn to controlled trading experiments in which the identity and objectives of the AI agents are known. Our baseline follows Smith et al. (1988), replacing human traders with autonomous AI agents. The setup of the experiment environment is as follows. Fundamental value is common knowledge and declines mechanically over a finite 15-period horizon. After the final dividend, the asset has no terminal value. Agents know the current period and exactly how many trading opportunities remain. We do not instruct them to speculate, coordinate, manipulate prices, or exploit other participants. Their objective is simply to maximize their own trading wealth.

We study two distinct agent architectures. LLM-based agents use a pretrained language model with persistent private memory and are not retrained on trading rewards during an experimental session. Reinforcement learning agents instead learn a trading policy directly from repeated market interaction and realized rewards using Proximal Policy Optimization (PPO).

The experimental results differ substantially across the two architectures. LLM-based agents initially trade actively but become increasingly conservative. Prices remain close to fundamental value (e.g., 59.7% of transactions occur exactly at fundamental value and only 1.9% occur above it). Unlike the human subjects in Smith et al. (1988), repeated interaction among the LLM agents does not generate persistent overpricing or a bubble pattern. Making the agents more forward-looking, introducing heterogeneous information, and allowing memories to persist across repetitions roughly doubles trading activity and increases mispricing, but the deviations still do not become self-reinforcing.

On the other hand, reinforcement learning agents produce a markedly different outcome. Trading behavior varies substantially across independently trained groups, with some groups generating much

larger price deviations and more aggressive trading than others. Approximately 80% of their transactions occur above fundamental value, yet only about 2% of the evaluation markets satisfy our operational bubble definition. We therefore find evidence that reinforcement learning can generate substantial speculative trading and mispricing, but no systematic evidence that the agents learn to influence or manipulate other traders, or generate persistent, self-reinforcing bubbles. We add several extensions to our trading experiments by changing the reward functions, increasing memory across repetitions, and giving agents more opportunities to trade and revise their quotes. These changes induce AI agents to trade more actively and generate greater asset overpricing, but the baseline agentic behavior remains unchanged.

The difference relative to Dou et al. (2025b) suggests that the emergence of destabilizing behavior is sensitive to market design. Our baseline asset has a transparent fundamental value, a known finite endpoint. By contrast, Dou et al. (2025b) study a long-lived asset over an infinite sequence of episodes with positive-feedback traders, constrained arbitrageurs, and intermediate periods in which prices can depart from fundamentals before realigning at the end of each episode. Their environment therefore creates a recurring intertemporal opportunity for planning agents to amplify temporary price distortions and profit from the induced feedback response. Our results suggest that autonomy and profit maximization alone are not sufficient to generate destabilizing behavior in a simpler environment with strong fundamental anchors. More generally, the comparison suggests that speculative trading and strategic exploitation are distinct outcomes, and that harmful AI trading depends on the interaction between agent capabilities and the opportunities created by market structure.

The paper makes three contributions. First, we study strategic exploitation as a distinct form of AI-driven behavior in financial markets. Calvano et al. (2020), Brown and MacKay (2023) and Cartea et al. (2026) show that algorithms can coordinate without communication and Dou et al. (2025a,b) show that planning agents can manipulate positive feedback traders. We ask more broadly when an individually profit-maximizing AI agent finds influencing another participant's behavior privately optimal, despite receiving no instruction to manipulate, coordinate, or cause harm. This research shifts attention from

whether a particular AI model is inherently destabilizing to the environments in which destabilizing behavior becomes profitable.

Second, we estimate an empirical measure of market susceptibility to feedback trading around a common information shock. Because it can be constructed directly from high-frequency data, it does not require observing trader identities or knowing whether individual market participants use AI.

Third, we conduct controlled AI agent experiments that reveal how autonomous agents with different architectures behave over time when placed in controlled market environments where their identities and objectives are known. Together, the two approaches separate the measurement of latent market vulnerability from the experimental analysis of whether AI agents generate strategic or destabilizing behavior. The financial stability consequences of autonomous trading systems therefore depend not only on agent capabilities, but also on the market structures, behavioral feedback effects, and intertemporal incentives with which those agents interact.

This paper proceeds as follows. Section 2 reviews the literature. Section 3 describes the data and the regression design. Section 4 presents the main results. Section 5 describes the experimental design and discusses various alterations to the experiment. Section 6 concludes.

## 2 Related Literature

The rapid spread of AI is changing how decisions are made in financial markets. For instance, Agrawal et al. (2018) and Cao et al. (2024) argue that AI lowers prediction costs but does not replace human judgment. In financial markets, where traders must interpret signals and respond to other traders, better prediction does not necessarily imply better prices or greater stability.

Agrawal et al. (2024) further argue that when decisions are interconnected, adopting AI in one area can require broader changes in workflows, coordination, and organizational design. Kazinnik and Brynjolfsson (2025) make a related point for central banks, where AI can improve information processing and operations but requires careful integration, data infrastructure, and oversight. Danielsson

and Uthemann (2025) and Lu et al. (2026) warn that widespread AI use may increase financial market fragility if institutions rely on similar data, models, and decision rules, because they may then react in similar ways at the same time.

Generative AI, and large language models (LLMs) in particular, create further risks because they make information, code, and analysis cheaper and faster to produce and apply (Brynjolfsson and Li, 2024). More recent literature studies AI trading agents (including LLMs) in theoretical settings and shows that they can learn to coordinate their actions without direct communication. Such behavior can reduce competition, weaken price efficiency, and create asset price bubbles (Calvano et al., 2020; Brown and MacKay, 2023; Dou et al., 2025a,b; Cartea et al., 2026).

Empirically, this outcome may not be as obvious as in theoretical settings because AI agents are trained on data produced by humans. As such, they may also replicate some human limitations. For instance, Horton et al. (2023) show that LLMs can reproduce some human-like behavior in economic experiments. Huang et al. (2026) show that when LLMs analyze the same financial dataset, their results are more similar than those of human researchers because generative AI tends to use similar methods.

Computer science studies further highlight the limitations of LLMs in real-world experiments. For instance, LLMs are not yet always able to maintain coherent decision-making over long periods (Backlund and Petersson, 2025), may fail to recognize poor terms in negotiation tasks (Zhang et al., 2026), and, most importantly for our setting, can be influenced by other LLMs in ways that lead to harmful outcomes (Wang et al., 2026).

# 3 Data and Methodology

## 3.1 Futures market data

We obtain intraday futures market data from LSEG Tick History for four underlying assets such as S&P 500 Index, crude oil, gold and bitcoin. The data cover the period from January 2019 through April 2026

and contain transaction records and best bid and offer quotes.

These futures contracts provide a useful setting for studying strategic AI behavior for two reasons. First, futures markets are highly automated. Evidence available in 2015 suggests that automated trading already accounted for between 60% and 80% of trading activity, meaning that the share may be even higher today (Hendershott and Riordan, 2013; Haynes and Roberts, 2015; Garcia-Ares et al., 2026). As such, prices, order flow, and market depth constitute the environment from which an AI trading agent would form its decisions. Second, the futures contracts include different underlying assets. Examining them permits relatively clean comparisons across asset classes to determine whether susceptibility to feedback trading is a general feature of automated markets or is concentrated in particular market structures.

## 3.2 Monetary policy events

We obtain FOMC event dates and event classifications from the Federal Reserve Bank of San Francisco's U.S. Monetary Policy Event-Study Database. The FOMC statements communicate the Committee's current policy decision and its contemporaneous assessment of economic conditions.

The San Francisco Fed database measures standard statement reactions over 30-minute windows. Our analysis instead uses substantially narrower windows because our objective differs from that of a conventional monetary policy event study. A 30-minute window is useful for measuring the total market response to a communication. It is too wide, however, for identifying whether an initial price movement induces subsequent trading. We therefore follow the high-frequency event-study literature and reconstruct the market responses from the underlying transaction and quotation records (see e.g., Ederington and Lee, 1993; Fleming and Remolona, 1999; Anderson et al., 2003; Bernile et al., 2016; Chordia et al., 2018).

## 3.3 Price responses

For each FOMC announcement and futures contract, we organize trades and quotations in event time. For FOMC event $t$, return horizon $h \in 1, 5, 10, 30, 60$ seconds, we define the post-FOMC log return as

$$r_{t,h} = \ln\left(\frac{P_{t,h-1}}{P_{t,-1}}\right), \tag{1}$$

where $P_{t,-1}$ is the final available price immediately before the FOMC announcement and $P_{t,h-1}$ is the final available price at the end of the return horizon. This procedure allows us to separate the immediate price response to public information from trading that occurs after the initial price movement.

## 3.4 Order flow

Our main measure is depth-scaled quote order flow imbalance (OFI). Following Cont et al. (2014), let $e_{t,n}$ denote the signed change in buying or selling pressure generated by quote update $n$ around FOMC event $t$. Positive values correspond to stronger displayed buying pressure, while negative values correspond to stronger displayed selling pressure. For order flow window $w \in \{5, 10, 30, 60, 120, 300, 600\}$ seconds, we sum these quote updates over $[h, h+w)$ and scale them by average displayed depth per side over the same window:

$$OFI^{\text{scaled}}_{t,h,w} = \frac{\sum_{n \in [h,h+w)} e_{t,n}}{\overline{D}_{t,h,w}}, \tag{2}$$

where $\overline{D}_{t,h,w}$ is average displayed depth per side. Scaling by depth preserves the direction of quote pressure while making its magnitude more comparable across markets with different levels of quoted liquidity.

For robustness, we use two alternative measures. First, unscaled quote OFI retains the original Cont et al. (2014) measure without the depth adjustment. Second, executed OFI captures directional pressure using completed trades. Each trade is matched to the most recent quotation observed strictly before the

transaction and classified as buyer- or seller-initiated. The two robustness measures are

$$OFI^{\text{quote}}_{t,h,w} = \sum_{n \in [h,h+w)} e_{t,n}, \tag{3a}$$

$$OFI^{\text{exec}}_{t,h,w} = \frac{\sum_{s=h}^{h+w-1} \left( BuyVolume_{t,s} - SellVolume_{t,s} \right)}{\sum_{s=h}^{h+w-1} TotalVolume_{t,s}}. \tag{3b}$$

Executed OFI lies between $-1$ and 1, with positive values indicating net buyer-initiated trading and negative values indicating net seller-initiated trading. Unlike the quote-based measures, it captures only completed transactions and not changes in displayed trading interest that occur without a trade.

## 3.5 Measure of market's susceptibility to feedback trading

This section develops an empirical proxy for a market's susceptibility to feedback trading. The approach is inspired by the theoretical study by Dou et al. (2025b), who develop a model of intertemporal coordination between AI agents.

A key implication of the model is that AI agents can learn to coordinate without explicit instruction or communication. They do so by creating and exploiting short-lived asset price bubbles, which may form and collapse within minutes after an information event. Whether this hidden coordination leads to market inefficiency depends on the mass of *feedback investors*, i.e., investors who buy after prices increase and sell after prices decrease, denoted by $\mu$. When $\mu$ is high, initial price movements attract more follow-on trading, which can amplify the original price movement and generate a bubble. When $\mu$ is low, this amplification mechanism is weaker, making bubbles less likely.

In the model, each information event can be divided into three sub-periods. In sub-period 0, the information event occurs. Sub-period 1 captures the immediate price reaction to the new information. Sub-period 2 captures the subsequent trading response of feedback investors. The parameter $\mu$ can therefore be interpreted as the sensitivity of feedback investors to earlier price movements. More

specifically, it measures how much order flow $y$ in sub-period 2 changes in response to the price change between sub-periods 0 and 1.

Building directly on this theory, our empirical strategy estimates the parameter $\mu$ using the following regression,

$$y_{t,2} = \mu(p_{t,1} - p_{t,0}) + \varepsilon_t \tag{4}$$

where $t$ indexes the information event. The dependent variable, $y_{t,2}$, is order flow in sub-period 2. The explanatory variable, $p_{t,1} - p_{t,0}$, is the price change between sub-periods 0 and 1. The term $\varepsilon_t$ captures shocks not directly explained by the model. In our empirical setting, a higher $\mu$ identifies a market environment in which prices created by earlier trades exert a stronger influence on subsequent trading decisions.

We estimate how $\mu$ varies with the market environment. In particular, we allow the susceptibility measure to depend on pre-announcement depth, order imbalance and recent trading activity. This approach is economically motivated because strategic trading behavior should be more profitable when an agent's orders have greater price impact and other participants respond more predictably to the resulting movement. Limited depth may increase price impact, while heightened volatility or complex information may make it more difficult for other participants to distinguish an informational price movement from a strategically generated signal.

We use these estimates to assign each contract-event observation a predicted feedback trading susceptibility measure. This procedure allows us to quantify the market environment immediately surrounding each announcement and identify the events during which feedback trading would have been most feasible.

## 3.6 Financial instability

In the final part of the analysis, we examine whether market environments with greater susceptibility to feedback trading (i.e., a proxy for when AI agents can behave strategically) are also associated with

greater financial instability. We proxy for financial instability using volatility amplification, i.e., the extent to which realized volatility after the feedback estimation interval exceeds its pre-announcement level. This measure captures whether market conditions associated with stronger feedback trading are followed by an increase in short-horizon price variability.

We first construct one-second prices from transaction data. One-second log returns, expressed in percentage points, are then defined as

$$r_{i,t,\tau} = 100\left[\ln(P_{i,t,\tau}) - \ln(P_{i,t,\tau-1})\right], \tag{5}$$

where $i$ indexes the contract, $t$ indexes the FOMC announcement, and $\tau$ denotes time in seconds relative to the announcement.

The timing of the instability measure is designed to separate the price movement used to estimate $\mu$, the subsequent order flow response, and the volatility outcome. For a return horizon $h$ and an order flow window $w$, the three intervals are

$$[0,h], \qquad [h,h+w], \qquad [h+w,h+w+900).$$

The first interval measures the initial post-announcement return, the second measures the subsequent order flow response used to estimate $\mu$, and the third is the 15-minute window over which post-event volatility is measured. Thus, the volatility window starts only after the complete feedback estimation interval has ended, preventing the instability measure from mechanically incorporating the price changes used to construct $\mu$. Pre-event volatility is measured over the common 30-minute window $[-1800,0)$.

For each event $t$, we calculate realized volatility as the sample standard deviation of one-second log returns within the corresponding pre- and post-event windows. Let $\sigma_{i,t}^{\text{pre}}$ denote volatility over $[-1800,0)$

and $\sigma^{\text{post}}_{i,t,h,w}$ denote volatility over $[h+w, h+w+900)$. We then define the event-specific volatility ratio as

$$V_{i,t,h,w} = \frac{\sigma^{\text{post}}_{i,t,h,w}}{\sigma^{\text{pre}}_{i,t}}. \tag{6}$$

A ratio greater than one, $V_{i,t,h,w} > 1$, indicates that realized volatility increased relative to the pre-announcement benchmark and therefore represents greater volatility amplification. A value equal to one indicates no change in volatility, while a value below one indicates that post-event volatility is lower than its pre-announcement level. We require at least 30 valid one-second return observations for each volatility estimate to be retained.

The volatility ratio is first calculated separately for every FOMC announcement. We then average these event-specific ratios across all valid announcements for each contract–return-horizon–order-flow-window combination:

$$\bar{V}_{i,h,w} = \frac{1}{N_{i,h,w}} \sum_{t=1}^{N_{i,h,w}} V_{i,t,h,w}, \tag{7}$$

where $N_{i,h,w}$ is the number of FOMC announcements with a valid event-specific volatility ratio for contract $i$ and the $(h,w)$ specification. This construction gives each valid FOMC event equal weight and produces one financial instability measure corresponding directly to each $\mu_{i,h,w}$ estimate used in the subsequent regressions.

# 4 Preliminary Empirical Results

This section examines the central empirical question of the paper, i.e., whether higher feedback trading susceptibility, as captured by $\widehat{\mu}$, is associated with greater subsequent financial instability. We measure financial instability using the post-event to pre-event volatility ratio developed in Section 3.

## 4.1 Descriptive Evidence

Table 1 reports summary statistics for the variables used in the empirical analysis. The post-event to pre-event volatility ratio averages 4.288 and has a median of 2.140, indicating substantial volatility amplification following FOMC announcements. There is also considerable variation in the feedback measure. For the depth-scaled quote-OFI specification, standardized $\widehat{\mu}$ has a mean of 0.037, a median of 0.065, and a standard deviation of 0.234, with estimates ranging from $-0.643$ to 0.594. Thus, the sample contains substantial variation in both measured feedback susceptibility and subsequent volatility.

Tables 2–4 provide a more detailed description of the two components used to estimate $\mu$. Table 2 shows that the dispersion of the initial post-announcement return generally increases with the return horizon, particularly for bitcoin and S&P 500 futures. Table 3 shows substantial variation in depth-scaled quote OFI across contracts and order flow windows, with dispersion generally increasing over longer windows. Table 4 reports the corresponding distributions for executed OFI and unscaled quote OFI. The alternative measures differ substantially in their units and dispersion. Executed OFI is bounded between $-1$ and 1, whereas unscaled quote OFI can take very large values and exhibits pronounced cross-market differences. These differences motivate the use of depth-scaled quote OFI as our preferred order flow measure and, subsequently, the use of standardized $\widehat{\mu}$ when comparing feedback across markets and measurement horizons.

The outcome variable also exhibits substantial cross-market variation. Table 5 shows average volatility ratios of approximately 1.6 for crude oil futures, around 2.0 for bitcoin, and roughly 2.1–2.3 for S&P 500 futures. Gold displays substantially greater volatility amplification, with average ratios generally above 10. The volatility ratio exceeds one for every contract–return-horizon–order-flow-window specification in the current sample. These pronounced differences across contracts motivate the inclusion of ticker fixed effects in the main regressions.

Overall, the descriptive evidence establishes meaningful variation in both sides of our empirical test, i.e., the return and order flow responses used to estimate $\mu$, and the subsequent volatility outcome used to measure financial instability.

## 4.2 Feedback trading

We next examine the estimated strength of feedback trading susceptibility. Tables 6–9 report $\widehat{\mu}$ for bitcoin, crude oil, S&P 500, and gold futures, respectively, using the depth-scaled quote OFI measure. The main feature of these estimates is their substantial heterogeneity across both markets and measurement horizons. For bitcoin, feedback is weak or negative in several very short-horizon specifications but becomes predominantly positive at longer return horizons. At the 60-second return horizon, for example, $\widehat{\mu}$ is positive across all order flow windows and statistically significant in most specifications. Crude oil futures display a different pattern. Feedback is positive in a number of short- and intermediate-horizon specifications, but becomes negative across several longer-horizon combinations.

The S&P 500 estimates are similarly heterogeneous, with the sign and magnitude of $\widehat{\mu}$ changing across return horizons and order flow windows. Gold displays particularly strong horizon dependence. At the five-second return horizon, the estimates are large and negative across all order flow windows, whereas at longer horizons they become substantially less negative and, in several specifications, positive. The broader implication is that market's susceptibility to feedback trading is not well characterized as a fixed property of an asset. Its measured strength depends materially on the horizon over which the initial price movement and subsequent trading response are observed.

Tables 10–13 repeat this analysis using executed OFI and unscaled quote OFI for the four contracts. The detailed coefficients differ from those obtained with the depth-scaled measure, but the same broad feature remains. The strength and sometimes the sign of estimated feedback depend on the market, horizon, and order flow definition. For example, executed OFI produces several positive and statistically significant coefficients for bitcoin at the five- and ten-second return horizons, while crude-oil estimates are generally smaller and mixed in sign. Executed OFI feedback is more consistently positive for the S&P 500 at shorter horizons, whereas the gold estimates again display substantial horizon dependence and are based on fewer usable transaction observations.

Taken together, these results establish substantial variation in feedback trading susceptibility across markets and horizons, while also showing that its measured magnitude is sensitive to the construction of

order flow.

## 4.3 $\widehat{\mu}$ and Financial Instability

We now turn to the central empirical test. Table 14 reports regressions of the post-event to pre-event volatility ratio on standardized $\widehat{\mu}$. If markets with stronger susceptibility to feedback trading exhibit higher financial instability, the coefficient on $\widehat{\mu}$ should be positive. As emphasized above, this prediction is a joint one. Greater feedback trading susceptibility must be accompanied by agents capable of recognizing and strategically exploiting that feedback for it to generate the destabilizing mechanism motivating the paper.

The first three columns use the preferred depth-scaled quote OFI measure. Column (1) reports pooled OLS. The coefficient on standardized $\widehat{\mu}$ is $-2.304$ and statistically significant at the 1% level. The unconditional relationship is therefore negative. Specifications with higher measures of feedback trading susceptibility are followed by lower, rather than higher, volatility amplification.

Adding ticker fixed effects in column (2) reduces the coefficient substantially in magnitude to $-0.634$, although it remains negative and statistically significant at the 1% level. Thus, a meaningful part of the unconditional relationship reflects persistent differences across the four contracts, but cross-contract heterogeneity does not account for it entirely. Column (3) additionally includes return-horizon and order flow window fixed effects. The coefficient falls further to $-0.281$ and is statistically significant at the 10% level. Using the cross-cell standard deviation of standardized quote-based $\widehat{\mu}$, 0.234, a one-standard-deviation increase in $\widehat{\mu}$ corresponds to a reduction of approximately 0.066 in the volatility ratio. Relative to the sample mean volatility ratio of 4.288, the estimated magnitude is modest, at approximately 1.5%.

The alternative order flow measures provide a useful qualification. When $\widehat{\mu}$ is estimated using unscaled quote OFI, the results closely resemble those for the preferred measure. The coefficients are $-1.658$ in pooled OLS, $-0.453$ with ticker fixed effects, and $-0.285$ in the fully controlled specification.

Most notably, the latter is almost identical to the corresponding depth-scaled estimate of $-0.281$. The negative relationship therefore does not appear to be driven by depth normalization. By contrast, when $\widehat{\mu}$ is estimated from executed OFI, the coefficients are positive—0.568, 0.753, and 0.526 across the three specifications—and only the ticker-fixed-effect estimate is marginally statistically significant. The negative relationship is therefore consistent across the two quote-based measures but is not reproduced when feedback is measured solely from completed transactions.

Taken at face value, the empirical analysis therefore does not support the hypothesis that higher $\mu$ is associated with greater subsequent financial instability in the current sample. For the preferred measure, the estimated association is instead modestly negative. This finding needs to be interpreted carefully, however, because the hypothesized positive relationship does not follow from a high $\mu$ alone.

In particular, the destabilizing mechanism requires sufficiently capable strategic agents to be present, to recognize that other traders respond predictably to earlier price movements, and to find it profitable to influence those movements and exploit the resulting order flow response. The historical data cannot tell us whether AI planning agents with these capabilities were operating in these markets during the sample period. More importantly, even if such agents were present, transaction and quotation records cannot reveal whether they were attempting to exploit feedback traders. The empirical hypothesis is therefore effectively a joint hypothesis about market susceptibility and its strategic activation.

The absence of a positive relationship between $\widehat{\mu}$ and volatility amplification may consequently mean that the measured feedback channels were not being systematically exploited in these securities during the period considered. It does not establish that sufficiently capable AI agents could not exploit such a channel, or that exploitation would not generate instability if it occurred. Nor can the historical data establish the stronger claim that no exploiting agents were present. The most that can be inferred from the observational analysis is that we find no evidence that greater measured feedback trading susceptibility was systematically associated with greater subsequent volatility amplification during this sample.

There is also a nonstrategic interpretation of the negative association. Around a genuine public information shock, a strong order flow response in the direction of the initial price movement may reflect

relatively rapid incorporation of the announcement into trading. If so, stronger measured feedback could coincide with faster information absorption and leave less unresolved price adjustment for the subsequent volatility window. The present empirical design cannot distinguish this interpretation from the possibility that potentially exploitable feedback simply remained unexploited. We therefore do not interpret the negative coefficient as evidence that positive-feedback trading is inherently stabilizing.

## 4.4 Robustness and Interpretation

Table 15 provides an additional robustness check using raw rather than standardized $\widehat{\mu}$. The results are less stable across order flow definitions. For depth-scaled quote OFI, the coefficient is negative and statistically significant in pooled OLS but becomes economically very small and statistically insignificant once fixed effects are introduced. A similar pattern arises for unscaled quote OFI. In contrast, raw $\widehat{\mu}$ estimated from executed OFI produces positive and statistically significant coefficients in the two fixed-effect specifications. Because raw $\widehat{\mu}$ inherits the units and dispersion of the underlying order flow measure, its scale differs substantially across markets and OFI definitions. The raw estimates are therefore less suitable for cross-market comparisons and reinforce our preference for standardized $\widehat{\mu}$ in the main specification.

The current results should also be interpreted in light of the relatively narrow empirical sample. The analysis presently covers four economically distinct futures contracts and FOMC announcements. This limits the extent to which we can distinguish a general relationship between feedback trading susceptibility and instability from features specific to the assets and events currently included.

Ongoing work expands the analysis to several hundred securities across a broader range of asset classes, return horizons, order flow windows, and information-event types. This substantially larger and more diverse sample will allow us to assess whether the relationship between $\mu$ and financial instability generalizes across market environments and will significantly strengthen the external validity of the results.

Finally, the identification limitation of the historical data provides the motivation for the experimental analysis in Section 5. The market analysis tells us whether a potentially exploitable feedback channel exists and whether greater measured susceptibility to feedback trading has historically coincided with greater instability. It cannot tell us whether AI agents were present, what objectives they pursued, or whether they were deliberately attempting to exploit the response captured by $\mu$. The experiments address precisely this missing part of the mechanism. The agents are known to us, their objectives and information are controlled, and their trading decisions and market consequences can be observed directly.

The two parts of the paper therefore answer complementary questions. The empirical analysis measures the market's susceptibility to feedback trading and asks whether that susceptibility has historically been associated with instability. The experiments ask whether autonomous AI agents can recognize and activate such vulnerabilities when they are actually present. This distinction is essential for interpreting the current market result. The absence of a positive unconditional relationship between $\mu$ and financial instability does not rule out the destabilizing mechanism. It shows instead that vulnerability alone is not sufficient; whether it becomes realized instability depends on the capabilities, incentives, and behavior of the agents operating in the market.

# 5 Experimental Design

We study the behavior of autonomous AI agents in a controlled experimental asset market. The market follows the design of Smith et al. (1988) but replaces human participants with independent AI agents. This setting allows us to examine whether AI agents trade at fundamental values or generate excessive trading (e.g., persistent mispricing, speculation, and exploitation) that can lead to bubbles and crashes.

Each experimental session contains nine traders, one risky asset, and 15 trading periods. At the end of every period, each share pays a common dividend of 0, 8, 28, or 60 cents, with equal probability. The asset has no terminal value after the final dividend. Its expected fundamental value is therefore known to all traders and declines mechanically as the number of remaining dividend payments falls. Three agents

begin with 225 cents and three shares, three with 585 cents and two shares, and three with 945 cents and one share. These endowments give all agents the same expected initial wealth but different combinations of cash and shares, ensuring that some traders have greater capacity to buy while others have greater capacity to sell.

Trading takes place in an order book where agents can submit buy and sell offers, and trades occur when a buy and a sell orders match. The highest bid and lowest ask receive priority, and a crossing order executes immediately against the standing quote. Agents cannot borrow money or short-sell shares. Furthermore, they control only their trading decisions and cannot alter the market rules.

The main purpose of this experimental design is to examine whether prices deviate from fundamental value. If they do, the deviation cannot easily be attributed to news, asymmetric information, or uncertainty about the asset's fundamental value. Instead, it must arise from the agents and their trading interactions. Because all agents within an experiment use the same model and receive the same general instructions, differences in behavior arise from their initial endowments, trading decisions, and private memories rather than from differences in model capability or training. To test different AI models, we also repeat the experiment using different AI agent specifications.

## 5.1 LLM-based Agents vs. Reinforcement Learning Agents

In general, we distinguish between LLM-based and non-LLM-based AI agents. One important class of non-LLM-based agents uses reinforcement learning (see, e.g., Dou et al., 2025a,b). In a trading setting, the agent observes a market state (e.g., prices, order book information, volatility, or past returns), chooses an action, and then receives a reward or penalty (profit or loss) based on the outcome. Over repeated interactions, the agent learns which actions maximize its objective. Its behavior is learned directly from numerical inputs, model architecture, and the reward function.

By contrast, LLM-based agents are built around pretrained language models. Although LLMs are also trained using reinforcement learning, these rewards generally reflect language quality, instruction

following, and human feedback rather than trading performance itself. In our experiment, LLM-based agents likewise do not receive qualitative market information. Their main advantage is that modern LLMs are among the most capable AI agent systems in the world and can make decisions without first being trained specifically for the experimental market. Their trading behavior therefore depends on the underlying LLM, the instructions they receive, the information available in their memory, and their repeated interactions with the market.

Finally, it is also important to distinguish LLM-based agents from isolated LLM calls (see, e.g., Cartea et al., 2026; Hansen and Lee, 2025). In an isolated LLM call, the researcher sends a prompt to the model and receives a response. Unless conversation history is explicitly maintained, the model does not automatically remember that it represents the same trader in the next round. The researcher must therefore store the trader's history, provide the relevant information in subsequent calls, validate the response, and repeatedly call the model. An LLM-based agent uses the same underlying LLM but adds a system for managing repeated decision-making and state. We use OpenAI's Agents SDK and assign each trader a separate agent and private memory. This memory automatically maintains the trader's previous interactions across decisions, while structured outputs ensure that responses follow the required trading format.[1]

## 5.2 LLM-based Agents

Each trader is implemented as a separate agent using the OpenAI Agents SDK. All agents use the same underlying LLM but have separate identities and memories. Each agent can therefore remember its own previous interactions but cannot access the private history of other agents. We repeat the experiment 10 times independently. Memory is reset between repetitions, so information from one repetition cannot carry over to the next.

[1] Other agent frameworks provide similar functionality. LangGraph provides detailed control over state and the sequence of decisions in an agent workflow. Microsoft AutoGen supports multi-agent systems in which agents can communicate and interact with one another. CrewAI allows agents to be assigned specific roles and tasks and organizes their interactions through workflows. These frameworks differ primarily in how they manage state, memory, communication, tools, and the sequence of agent interactions. They can use the same underlying language model, although the choice of framework can still affect observed behavior by changing how the model interacts with its environment.

Whenever an agent is activated, it receives information about its current cash and share holdings, the current and remaining periods, the dividend distribution, the best bid and ask, the most recent transaction price, recent trades, outcomes from the previous period, and its own recent actions. Each agent is instructed to maximize its final cash balance using only the information available to it. Direct communication and explicit coordination between agents are prohibited. The agent then returns only its selected action.[2]

In the original experimental design of Smith et al. (1988), participants could repeatedly submit bids and asks and react to other traders' orders during a trading period. In our implementation, each of the nine agents is randomly activated eight times within each of the 15 periods. This design gives the agents repeated opportunities to revise their quotes or accept new orders. Persistent memory thereby allows the agent to use information from its previous decisions and outcomes when making a new decision. For example, it may change its quotes after repeatedly failing to trade or after experiencing poor returns. It is important to note that the underlying LLM is not retrained during the experiment. Any change in behavior therefore reflects adaptation through remembered information rather than learning through changes in the model itself.

At the end of each period, we record all submitted actions, executed trades, transaction prices, trading volume, dividends, cash holdings, and share holdings, and document agent-level performance. Agent wealth is measured as cash plus the expected value of its remaining shares.

#### 5.2.1 Result #1: LLM-based agents do not speculate

We find that LLM-based agents actively participate in the market at the beginning of the experiment, but become more conservative over time. As shown in Figure 1, the agents execute an average of 4.8 trades in the first period and 3.6 trades in the second period. From period 7 onward, the median number of transactions per experimental repetition is zero.

[2]We do not ask the agent to explain its decision because a model-generated explanation does not reliably reveal the model's actual internal decision process.

**[Figure 1 about here]**

Trading activity also varies substantially across experimental repetitions. Across a full 15-period experiment, an average of 21.6 transactions are executed in each experimental repetition. The total number of transactions ranges from only 12 to 37 across the ten independent repetitions. Summary statistics are provided in Table 16.

**[Table 16 about here]**

Transaction prices remain remarkably close to fundamental value. Across the ten experimental repetitions, the average relative deviation from fundamental value is only $-0.315\%$, indicating a slight tendency toward underpricing. Of the 216 transactions executed across all ten experiments, 129 (59.7%) occur exactly at fundamental value. Another 83 transactions (38.4%) occur below fundamental value. Only four transactions (1.9%) of all trades occur above fundamental value. Similarly, among 938 submitted bids, only two are above fundamental value. Thus, the agents almost never express a willingness to purchase the asset at a price exceeding its expected remaining dividend value. Hence, repeated interaction between the LLM-based agents does not generate persistent overpricing or speculative dynamics. This behavior differs substantially from the original experiments of Smith et al. (1988). With human participants, 14 of the 22 experimental markets exhibited price bubbles followed by crashes relative to fundamental value.

The average cumulative profit at the end of the experiment is 64.8 cents per agent (i.e., average cumulative return of approximately 5%). We also compare each agent with a passive buy-and-hold strategy. The initial endowment thereby affects how agents benefit from trading. Agents in the third group (with 945 cents and one share) perform best in seven of the ten experimental repetitions. They earn an average excess profit of 31 cents relative to buy-and-hold. In contrast, agents in the first group (starting with 225 cents and three shares) earn an average excess profit of $-7.8$ cents, while agents in the second group (starting with 585 cents and two shares) earn $-23.2$ cents. The reason is that agents that start with more cash are able to buy more shares, while agents that start with more shares tend to

sell more. Because realized dividends are, on average, higher than expected, agents that buy additional shares tend to benefit more.

These findings provide an important benchmark for the following experiments, as speculative dynamics do not emerge in the baseline LLM setting. We therefore examine whether speculative trading emerges under different model specifications and information settings, for example when different agents are given less information, have more opportunities to interact, or place greater weight on future resale opportunities.

### 5.2.2 Alternative Specifications with LLM-based Agents

The first alternative specification is to add a forward-looking component to the agents' instructions. Agents continue to maximize their final cash balance, but they are also asked to consider future resale opportunities in their objective functions.

The second specification introduces differences in the information available to traders. In the baseline experiment, all agents receive the same information about the asset. Naturally, this design limits disagreement about the value. We therefore introduce three information groups. Full-information agents receive the same information as in the baseline experiment. Structural-information agents receive the dividend distribution and the remaining number of periods. Finally, market-information agents observe market prices, bids and asks, previous transactions, realized dividends, and the remaining trading horizon, but do not receive the dividend distribution or the fundamental value. Moreover, each of the three initial endowment groups contains one agent from each information group. This ensures that differences in information are not mechanically related to differences in initial cash or share holdings.

In the third and final specification, we relax the memory restriction and allow the same agents to retain their private memories across experimental repetitions. After each 15-period experiment, the market is reset with new initial endowments, an empty order book, and new dividends. However, the agents retain information from their previous trading experiences. They can therefore adapt their behavior based on what was previously successful or unsuccessful. Importantly, the underlying LLM is not retrained,

and its model parameters remain fixed. Any behavioral change therefore reflects adaptation through accumulated memory rather than changes in the model itself.

#### 5.2.3 Result #2: LLM-based agents change trading strategies

Changing the LLM-based agent specification increases trading activity and price deviations from fundamental value. Across the ten experimental repetitions, agents execute an average of 45.3 trades (i.e., trading activity more than doubled). Average mispricing also increases to 0.78%, indicating that prices are now slightly above fundamental value rather than slightly below it. Nonetheless, there is substantial variation across repetitions (min: $-7.22\%$, max: $4.42\%$). 243 of the 453 trades (about 54%) occurred above fundamental value. Fully informed traders earned an average excess profit of approximately \$0.73, structurally informed traders earned \$0.30, and market-only traders lost approximately \$1.03 relative to buy-and-hold.

There are at least two plausible reasons for this behavior. First, AI agents observe two conflicting signals, i.e., prices can increase, while fundamental value declines. Unlike human traders, a powerful pretrained LLM may recognize that buying above fundamental value is risky without a convincing resale opportunity. In contrast, Haruvy et al. (2007) show that human traders often extrapolate recent price increases and expect them to continue. LLM-based agents appear less willing to follow such trends when they conflict with the known declining fundamental value and finite horizon. Second, LLM-based agents do not automatically learn profitable trading rules from rewards during the experiment.

We also observe that agents follow three broad trading strategies. The first is passive behavior. In the most passive groups, approximately 68% of all decisions were hold decisions, and relatively few trades were completed. In the least active groups, only about two or three of the nine agents were active on average. In other groups, many agents submitted buy orders, but relatively few of these orders could be executed immediately. These agents appeared willing to buy only at selected prices rather than accept the best available selling price. Some of these groups produced relatively small differences in wealth among the agents.

Other experimental runs show the emergence of aggressive traders. These traders held less frequently, submitted more orders that could be executed immediately, and completed more trades. They appeared to prioritize completing trades quickly rather than waiting for more favorable prices. However, the group's average excess return relative to buy-and-hold remained approximately zero. The strategy therefore changed how wealth was distributed among the agents, but it did not generate an average advantage for the group as a whole. Finally, we observe a third strategy in which agents submit both buy and sell orders.

The strategies or roles adopted by traders were related to their access to information. For example, the most aggressive trader had access only to market information. Fully informed traders were less frequently aggressive and more frequently passive. However, they also behaved aggressively in some periods. Trading on both sides of the order book was observed among all three information types.

The traders' roles also changed substantially from period 1 to period 15. Across the experimental repetitions, none of the 90 trader histories (i.e., nine agents in each of 10 repetitions) followed the same strategy during all 15 periods. On average, a trader changed its behavioral role approximately seven times. During periods 1-5, approximately 10% of traders were classified as passive and 32% as aggressive. During periods 11-15, approximately 45% were passive and only 21% were aggressive. Trading on both sides of the order book also decreased, from approximately 24% during the first five periods to 14% during the final five periods.

## 5.3 Reinforcement Learning Agents

As an alternative to LLM-based agents, we implement trading agents that learn directly from market interaction through reinforcement learning. These type of agents are related to Dou et al. (2025b), who also study whether reinforcement learning agents can generate destabilizing trading behavior. However, their agents and ours differ importantly in how trading strategies are learned and represented. For instance, Dou et al. (2025b) use Q-learning, which, in its simplest form, assigns a value to each possible

combination of a market state and an action.[3] Furthermore, these Q-learning agents have an explicit planning ability, which allows agents to consider how a trading decision today may affect future prices, market conditions, and profits. They therefore ask whether agents that can plan ahead learn coordinated strategies that create and later exploit price distortions.

Our objective is different. We do not impose a particular planning or coordination structure on the agents. Instead, we ask whether speculative, coordinated, or strategic trading behavior emerges when AI agents repeatedly interact in the experimental market and learn from this interaction. We therefore use Proximal Policy Optimization (PPO) with a neural network rather than tabular Q-learning.

PPO directly learns how the agent should act in different market states. Instead of storing a separate value for every possible state-action combination, the neural network receives the current numerical market state and produces a trading decision. The inputs include the trader's cash and shareholdings, the remaining periods, the best bid and ask, recent transaction prices, previous market activity, and the trader's own recent actions. The output determines whether the trader holds, submits or accepts an order, cancels an order, and, where necessary, chooses a price.

A neural network is particularly useful in our setting because the Smith et al. (1988) market has a substantially richer state space than a stylized environment. A tabular Q-learning implementation would require us to divide variables into discrete categories and store a value for every relevant combination. The number of possible states then increases extremely quickly.[4] The neural network avoids this problem through function approximation and learns relationships across states.

### 5.3.1 Training

All nine traders share the same market environment and action space, but each trader maintains its own parameterized neural network and learns independently from its own observations and rewards.

[3] An agent can learn that buying is profitable when the price is low and selling is profitable when the price is high. After repeatedly experiencing states, choosing actions, and receiving rewards, the agent updates these values and increasingly chooses the actions that have generated the highest expected future payoff.

[4] Even with only ten state variables and ten possible values for each variable, the resulting table could contain $10^{10}$ states before accounting for the different possible actions. Most of these states would rarely or never be observed during training.

During training, the agents repeatedly trade in independently simulated versions of the experimental market. After observing the consequences of their decisions, PPO changes the weights of the neural network in directions that increase expected future rewards. The nine traders are trained simultaneously through repeated interaction in the same market. This design allows the behavior of one agent to affect the learning environment of the other agents and therefore permits strategic responses to emerge endogenously.

Each training episode consists of a complete 15-period experimental asset market. Traders choose among holding, canceling an outstanding order, accepting the best bid or ask, or submitting a new bid or ask. The price associated with a new order is chosen by the policy as a continuous action.[5] Like the setup with the LLM-based agents, the reinforcement learning agents observe their own cash and asset holdings, the remaining trading horizon, current and recent bids and asks, transaction prices, order-book depth, the bid-ask spread, recent public order-book states, realized dividends, and their own recent actions. Furthermore, agents must learn the economic implications of the dividend process and market history themselves. The trained agents can trade up to eight times per period, remember earlier market activity through a Long Short-Term Memory (LSTM) memory layer, and observe the eight most recent order-book states. The recurrent component allows agents to condition current actions on sequences of previous market states rather than treating each decision as independent.

As we do not want to directly reward speculation, coordination, or harming other traders, agents receive no reward during the market and are rewarded only at the end of period 15. The reward depends on whether trading improves final wealth relative to buy-and-hold. For trader $i$, the terminal reward is

$$R_i = \frac{W_{i,T} - W_{i,T}^{BH}}{100}, \tag{8}$$

where $W_{i,T}$ denotes final wealth and $W_{i,T}^{BH}$ denotes the wealth obtained by passively retaining the trader's initial endowment under the same realized sequence of dividends. Consequently, buying above

[5] To avoid learning from mechanically infeasible choices, the final specification uses action masking so that actions that cannot be executed because of insufficient cash, insufficient inventory, or the absence of an executable quote are excluded from the available action set.

fundamental value is not immediately punished if the asset can later be resold at a higher price. This design allows speculative trading to emerge without explicitly encouraging it.

We train ten independent groups of nine reinforcement learning agents. Training proceeds for 500 PPO iterations. To encourage agents to try different strategies early in training, we initially allow more exploration.[6] Agents therefore explore different strategies initially and increasingly rely on previously successful strategies as training progresses. Finally, we evaluate trading performance out-of-sample.

#### 5.3.2 Result #3: Reinforcement learning agents learn to speculate

We evaluate learning in markets that are separate from those used to train the neural networks. The analysis includes 10 independently trained groups, with nine reinforcement learning traders in each group. Each trained group was evaluated in 100 new experimental repetitions. One experimental repetition consists of a complete market with 15 trading periods.

Looking at average behavior per trader, we find that traders were generally more active at the beginning. Many became more cautious and passive toward the end. Once traders became passive, they were also more likely to remain passive in the following period. Because agents maximize final wealth rather than trading activity, this pattern is consistent with agents learning that many trades do not improve their final wealth. The reinforcement learning agents completed an average of 8.84 trades in each experimental repetition. The median across the independently trained groups was only 4.77 trades, however, and the group averages ranged from 1.81 to 31.44 trades.

At the same time, trading behavior differs across the independently trained groups of nine reinforcement learning agents.[7] For instance, some independently trained groups traded very little, while others completed more than 30 trades per market. Some learned traders almost always held, while others remained continuously active. The groups also differed substantially in their price deviations and

[6] The entropy coefficient is gradually reduced from 0.03 at the beginning of training to 0.01 halfway through training and 0.001 at the end.

[7] Importantly, these descriptions refer to the average behavior of each group. They do not mean that all nine agents in a group followed exactly the same strategy.

in the distribution of wealth among traders. This heterogeneity arises because all nine agents learn simultaneously. When one trader changes its strategy, it changes the market faced by the other traders, which may then adjust their own strategies. As a result, different groups can learn different trading strategies even though they operate under exactly the same market rules. Another possibility is that the learning process may have remained unstable, so the final strategies may not represent fully converged behavior. The present evaluation cannot completely distinguish between these explanations.

Across all decisions, 43.71% were hold decisions. At the trader level, some learned traders therefore almost never selected hold, whereas others became completely passive in the out-of-sample markets. Most completed transactions were initiated by buyers. Buyer-initiated trades accounted for 63.38% of all transactions. An average experimental repetition contained 6.91 aggressive buy executions and 1.94 aggressive sell executions. Furthermore, agents produced substantial deviations between transaction prices and fundamental value. The average relative deviation was 213.96%. Approximately 80.41% of all transactions occurred above fundamental value. However, only approximately 2% of the evaluation markets satisfied the operational price-bubble definition. This indicator requires prices to remain more than 5% above fundamental value for at least two consecutive periods and then decline by at least 5% after the price peak.

The distribution of profits was closely related to the initial endowments. Traders beginning with 225 cents and three shares earned an average excess profit of 111.48 cents by the end of period 1. Their advantage increased gradually and reached 158.42 cents by period 15. Traders beginning with 945 cents and one share experienced the opposite outcome. Their average excess profit was $-122.73$ cents after period 1 and declined to $-155.66$ cents by period 15. Traders beginning with 585 cents and two shares occupied an intermediate position. Their average excess profit was 11.25 cents in period 1 and approximately $-2.75$ cents by period 15.

# 6 Conclusion

This paper studies whether autonomous AI agents can learn harmful, exploitative, or collusive trading behavior and whether such behavior can threaten financial stability. We combine empirical regression analysis with controlled trading experiments using both LLM-based and reinforcement-learning agents.

In the empirical analysis, we find no evidence that greater market feedback vulnerability is associated with economically meaningful deterioration in financial stability in our current sample. This result does not imply that such feedback channels cannot be exploited. The historical data identify potentially exploitable market conditions but cannot reveal whether agents capable of strategically exploiting them were present or attempting to do so.

The experiments provide a test in a setting where the agents, their objectives, and their actions are directly observed. We find no evidence that strategic or harmful behavior emerges endogenously. LLM-based agents trade actively but remain close to fundamental value, and changes in their specification increase trading and price deviations without producing persistent self-reinforcing speculation. Reinforcement-learning agents produce substantially greater heterogeneity in trading strategies and price deviations. Some groups trade aggressively and generate larger wealth transfers, while others become increasingly passive. Despite these differences, we find no systematic evidence that the agents learn to coordinate, strategically exploit other traders, or generate persistent self-reinforcing bubbles.

Overall, the results suggest that autonomous AI trading can generate substantial trading activity, mispricing, and redistribution across market participants without necessarily producing strategic exploitation or financial instability. Across the market environments and agent designs considered here, we find no systematic evidence that market feedback vulnerability or endogenous AI trading behavior translates into economically meaningful deterioration in financial stability.

# References


Agrawal, A., Gans, J. S., and Goldfarb, A. (2018). Human judgment and AI pricing. In *AEA Papers and Proceedings*, volume 108, pages 58–63. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Agrawal, A., Gans, J. S., and Goldfarb, A. (2024). Artificial intelligence adoption and system-wide change. *Journal of Economics & Management Strategy*, 33(2):327–337.

Anderson, T. G., Bollerslev, T., Diebold, F. X., and Vega, C. (2003). Micro effects of macro announcements: Real-time price discovery in foreign exchange. *American Economic Review*, 93(1):38–62.

Backlund, A. and Petersson, L. (2025). Vending-bench: A benchmark for long-term coherence of autonomous agents. Working paper.

Bernile, G., Hu, J., and Tang, Y. (2016). Can information be locked up? informed trading ahead of macro-news announcements. *Journal of Financial Economics*, 121(3):496–520.

Brown, Z. Y. and MacKay, A. (2023). Competition in pricing algorithms. *American Economic Journal: Microeconomics*, 15(2):109–156.

Brynjolfsson, E. and Li, D. (2024). The economics of generative AI. *NBER Reporter*, (1):16–19.

Calvano, E., Calzolari, G., Denicolo, V., and Pastorello, S. (2020). Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review*, 110(10):3267–3297.

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

Cartea, A., Chang, P., Chen, N., and Zhong, M. (2026). AI bubbles with large language models. Working paper.

Chordia, T., Green, T. C., and Kottimukkalur, B. (2018). Rent seeking by low-latency traders: Evidence from trading on macroeconomic announcements. *Review of Financial Studies*, 31(12):4650–4687.

Cont, R., Kukanov, A., and Stoikov, S. (2014). The price impact of order book events. *Journal of Financial Econometrics*, 12(1):47–88.

Danielsson, J. and Uthemann, A. (2025). Artificial intelligence and financial crises. *Journal of Financial Stability*, page 101453.

Dou, W. W., Goldstein, I., and Ji, Y. (2025a). AI-powered trading, algorithmic collusion, and price efficiency. Working paper.

Dou, W. W., Goldstein, I., and Ji, Y. (2025b). Financial market fragility in the era of AI planning. Working paper.

Ederington, L. H. and Lee, J. H. (1993). How markets process information: News releases and volatility. *Journal of Finance*, 48(4):1161–1191.

Fleming, M. J. and Remolona, E. M. (1999). Price formation and liquidity in the us treasury market: The response to public information. *Journal of Finance*, 54(5):1901–1915.

Garcia-Ares, P. A., Amaya, D., Pearson, N. D., and Vasquez, A. (2026). The rise of algorithmic retail option traders. Working paper.

Hansen, A. L. and Lee, S. J. (2025). Financial stability implications of generative AI: Taming the animal spirits. Working paper.

Haruvy, E., Lahav, Y., and Noussair, C. N. (2007). Traders' expectations in asset markets: experimental evidence. *American Economic Review*, 97(5):1901–1920.

Haynes, R. and Roberts, J. S. (2015). Automated trading in futures markets. *CFTC White Paper*, 15:19.

Hendershott, T. and Riordan, R. (2013). Algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis*, 48(4):1001–1024.

Horton, J. J., Filippas, A., and Manning, B. S. (2023). Large language models as simulated economic agents: What can we learn from homo silicus? Working paper.

Huang, W., Menkveld, A. J., and Yu, S. (2026). AI "errors". Working paper.

Kazinnik, S. and Brynjolfsson, E. (2025). AI and the fed. Working paper.

Lu, R., Xu, Y., and Vulicevic, L. (2026). Do LLMs make markets more efficient? Working paper.

MacDiarmid, M., Wright, B., Uesato, J., Benton, J., Kutasov, J., Price, S., Bouscal, N., Bowman, S., Bricken, T., Cloud, A., et al. (2025). Natural emergent misalignment from reward hacking in production rl. *arXiv preprint arXiv:2511.18397*.

Smith, V. L., Suchanek, G. L., and Williams, A. W. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica*, pages 1119–1151.

Wang, S., Politi, M., Marro, S., and Crapis, D. (2026). Profit is the red team: Stress-testing agents in strategic economic interactions. Working paper.

Zhang, E., Zhang, F., Pappu, A., El, B., Blanchet, J., Athey, S., Liu, J., and Zou, J. (2026). TERMS-bench: Diagnosing LLM negotiation agents beyond deal rate. Working paper.

# Tables and Figures

**Table 1. Descriptive Statistics**

This table pools the variables used in the empirical analysis across the relevant tickers and horizon specifications. Post-event log returns are expressed in percentage points. Depth-scaled quote OFI is the main order-flow measure; executed OFI is reported as a robustness measure. Raw and standardized $\mu$ statistics are reported for executed OFI and depth-scaled quote OFI. The volatility ratio is first calculated separately for each FOMC announcement as post-event volatility divided by pre-event volatility and is then averaged across valid events for each ticker–horizon–window specification. *N* is the number of non-missing observations contributing to each variable and therefore need not be identical across rows.

| Variable | *N* | Mean | Median | SD | Min | Max |
|---|---|---|---|---|---|---|
| Post-event log return (%) | 1060 | 0.0059 | 0 | 0.2869 | -3.6815 | 2.2118 |
| Executed OFI | 6734 | 0.0148 | 0.0227 | 0.3838 | -1.0000 | 1.0000 |
| Depth-scaled quote OFI | 8644 | 12.60 | 3.6205 | 160.01 | -1,436.8 | 2,550.1 |
| Raw mu, executed OFI | 113 | 0.0646 | 0.0723 | 0.7582 | -3.1740 | 3.6055 |
| Standardized mu, executed OFI | 113 | 0.0727 | 0.0678 | 0.1322 | -0.3582 | 0.4398 |
| Raw mu, quote OFI | 133 | -20.92 | 17.01 | 317.15 | -1,426.4 | 692.83 |
| Standardized mu, quote OFI | 133 | 0.0368 | 0.0653 | 0.2335 | -0.6425 | 0.5937 |
| Volatility ratio | 140 | 4.2880 | 2.1399 | 4.3047 | 1.4717 | 13.99 |

**Table 2. Post-Event Returns by Ticker and Return Horizon**

Post-event returns are log returns expressed in percentage points and are measured from the FOMC announcement to return horizon $h$. $N$ is the number of FOMC announcements for which the prices required to construct a valid return are available for that ticker and horizon. Consequently, $N$ can differ across tickers and horizons when individual events lack usable price observations.

| Ticker | $h$ | $N$ | Mean | Median | SD | Min | Max |
|---|---|---|---|---|---|---|---|
| CLc1 | 1 | 64 | 0.0032 | 0.0005 | 0.0242 | -0.0620 | 0.0491 |
| CLc1 | 5 | 64 | 0.0036 | 0 | 0.1085 | -0.2937 | 0.5483 |
| CLc1 | 10 | 64 | 0.0206 | 0.0009 | 0.1530 | -0.4531 | 0.7237 |
| CLc1 | 30 | 64 | 0.0497 | 0.0283 | 0.1998 | -0.5985 | 0.9552 |
| CLc1 | 60 | 64 | 0.0729 | 0.0206 | 0.2913 | -0.6804 | 1.2972 |
| GCc1 | 1 | 21 | 0 | 0 | 0 | 0 | 0 |
| GCc1 | 5 | 21 | 0.0070 | 0 | 0.0601 | -0.1074 | 0.2486 |
| GCc1 | 10 | 21 | 0.0109 | 0 | 0.0991 | -0.2055 | 0.2890 |
| GCc1 | 30 | 21 | -0.0089 | 0 | 0.1646 | -0.5320 | 0.3179 |
| GCc1 | 60 | 21 | 0.0070 | 0 | 0.1354 | -0.2598 | 0.3556 |
| ESc1 | 1 | 63 | -0.0116 | 0 | 0.0489 | -0.1844 | 0.0751 |
| ESc1 | 5 | 63 | -0.0274 | -0.0037 | 0.1620 | -0.6323 | 0.3468 |
| ESc1 | 10 | 63 | -0.0001 | 0.0020 | 0.2236 | -0.5680 | 0.8159 |
| ESc1 | 30 | 63 | 0.0340 | 0.0394 | 0.3071 | -0.8136 | 1.7016 |
| ESc1 | 60 | 63 | 0.0698 | 0.0614 | 0.3945 | -0.9800 | 2.2118 |
| BTCc1 | 1 | 64 | -0.0097 | 0 | 0.0354 | -0.1878 | 0.0488 |
| BTCc1 | 5 | 64 | -0.0412 | 0 | 0.2089 | -1.0370 | 0.3746 |
| BTCc1 | 10 | 64 | -0.0500 | 0 | 0.2921 | -1.0570 | 0.6290 |
| BTCc1 | 30 | 64 | -0.0062 | 0 | 0.5346 | -2.4792 | 1.0141 |
| BTCc1 | 60 | 64 | -0.0143 | 0 | 0.6673 | -3.6815 | 1.0298 |

**Table 3. Depth-Scaled Quote OFI by Ticker and Order-Flow Window**

Quote OFI is constructed following Cont, Kukanov, and Stoikov (2014) from signed best-bid and best-ask update contributions and is scaled by average displayed depth per side over the same order-flow window. Statistics are reported by ticker and order-flow window $w$, pooling across the return horizons used in the analysis. $N$ counts non-missing event–return-horizon observations contributing to each ticker–window cell and is therefore not the number of distinct FOMC announcements.

| Ticker | $w$ | $N$ | Mean | Median | SD | Min | Max |
|---|---|---|---|---|---|---|---|
| CLc1 | 5 | 320 | 0.6981 | 0.1713 | 13.74 | -93.75 | 68.08 |
| CLc1 | 10 | 320 | 2.1754 | 1.4669 | 19.01 | -114.29 | 122.55 |
| CLc1 | 30 | 320 | 5.2127 | 3.5492 | 25.66 | -103.28 | 122.63 |
| CLc1 | 60 | 320 | 3.1266 | 2.2049 | 42.98 | -253.67 | 168.45 |
| CLc1 | 120 | 320 | 0.8039 | 3.9700 | 52.19 | -254.52 | 181.99 |
| CLc1 | 300 | 320 | -3.8368 | 8.5401 | 94.86 | -391.74 | 320.07 |
| CLc1 | 600 | 325 | -10.78 | 2.3082 | 119.79 | -421.66 | 361.92 |
| GCc1 | 5 | 275 | 3.3608 | 2.8571 | 32.41 | -140.67 | 139.00 |
| GCc1 | 10 | 275 | 1.8958 | 4.4898 | 45.27 | -197.58 | 154.00 |
| GCc1 | 30 | 275 | 5.4151 | 7.4510 | 76.13 | -336.24 | 292.43 |
| GCc1 | 60 | 275 | 10.54 | 11.48 | 99.17 | -335.52 | 749.58 |
| GCc1 | 120 | 275 | 31.73 | 18.42 | 154.57 | -323.95 | 1,016.6 |
| GCc1 | 300 | 275 | 65.96 | 43.16 | 241.24 | -312.94 | 1,632.3 |
| GCc1 | 600 | 275 | 72.71 | 41.38 | 299.46 | -567.88 | 1,830.6 |
| ESc1 | 5 | 317 | 5.4464 | 3.4152 | 45.10 | -327.89 | 191.36 |
| ESc1 | 10 | 317 | 7.5028 | 3.4532 | 54.71 | -248.84 | 246.82 |
| ESc1 | 30 | 320 | 20.64 | 18.85 | 73.75 | -276.71 | 315.40 |
| ESc1 | 60 | 320 | 23.65 | 17.32 | 99.30 | -296.35 | 436.55 |
| ESc1 | 120 | 320 | 34.65 | 22.57 | 140.28 | -305.33 | 722.44 |
| ESc1 | 300 | 320 | 47.36 | 24.55 | 239.60 | -462.78 | 1,554.3 |
| ESc1 | 600 | 320 | 42.21 | 27.57 | 374.95 | -720.82 | 2,550.1 |
| BTCc1 | 5 | 320 | 1.8123 | 0 | 44.09 | -209.23 | 490.00 |
| BTCc1 | 10 | 320 | 2.4203 | 0 | 56.56 | -297.60 | 338.26 |
| BTCc1 | 30 | 320 | 0.8141 | 0 | 99.39 | -426.17 | 534.29 |
| BTCc1 | 60 | 320 | -6.9321 | 0 | 135.90 | -680.00 | 629.21 |
| BTCc1 | 120 | 320 | -8.5913 | 1.5105 | 178.14 | -765.75 | 663.92 |
| BTCc1 | 300 | 320 | -9.5047 | 5.5620 | 272.03 | -1,167.3 | 885.32 |
| BTCc1 | 600 | 320 | 17.23 | 13.59 | 333.92 | -1,436.8 | 1,046.6 |

**Table 4. Descriptive Statistics for Robustness Order-Flow Measures**

The table reports the two robustness order-flow measures side-by-side for each ticker and order-flow window $w$. Executed OFI is defined as net signed executed volume divided by total executed volume. Unscaled quote OFI is constructed following Cont, Kukanov, and Stoikov (2014) from signed best-bid and best-ask update contributions. Statistics are reported by ticker and order-flow window, pooling across the return horizons used in the analysis. $N$ counts non-missing event–return-horizon observations contributing to each ticker–window cell and is therefore not the number of distinct FOMC announcements. Reported statistics are $N$, mean, median, standard deviation, minimum, and maximum.

| | | Executed OFI | | | | | | Unscaled Quote OFI | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ticker | $w$ | $N$ | Mean | Median | SD | Min | Max | $N$ | Mean | Median | SD | Min | Max |
| CLc1 | 5 | 310 | 0.0335 | 0.0640 | 0.5238 | -1.0000 | 1.0000 | 325 | 6.6246 | 1.0000 | 82.87 | -389.00 | 758.00 |
| CLc1 | 10 | 317 | 0.0386 | 0.0833 | 0.4400 | -1.0000 | 1.0000 | 325 | 13.90 | 9.0000 | 105.16 | -343.00 | 589.00 |
| CLc1 | 30 | 320 | 0.0247 | 0.0228 | 0.2930 | -1.0000 | 0.9091 | 325 | 40.38 | 25.00 | 182.03 | -587.00 | 885.00 |
| CLc1 | 60 | 320 | 0.0313 | 0.0186 | 0.2105 | -0.4388 | 0.8261 | 325 | 38.17 | 15.00 | 283.63 | -1,410.0 | 1,181.0 |
| CLc1 | 120 | 320 | 0.0302 | 0.0232 | 0.1741 | -0.3430 | 0.8235 | 325 | 21.73 | 45.00 | 350.78 | -1,402.0 | 1,186.0 |
| CLc1 | 300 | 320 | 0.0260 | 0.0178 | 0.1483 | -0.3127 | 0.7442 | 325 | 65.58 | 83.00 | 607.73 | -2,088.0 | 2,161.0 |
| CLc1 | 600 | 320 | 0.0070 | 0.0106 | 0.1241 | -0.3739 | 0.5127 | 325 | 7.3138 | 39.00 | 840.40 | -2,111.0 | 2,372.0 |
| GCc1 | 5 | 11 | -0.0118 | 0.2000 | 0.9326 | -1.0000 | 1.0000 | 295 | 7.9898 | 3.0000 | 84.63 | -385.00 | 364.00 |
| GCc1 | 10 | 16 | 0.0327 | 0.2923 | 0.9348 | -1.0000 | 1.0000 | 295 | 3.1763 | 4.0000 | 119.64 | -717.00 | 398.00 |
| GCc1 | 30 | 25 | -0.1772 | 0 | 0.7910 | -1.0000 | 1.0000 | 295 | 8.6068 | 9.0000 | 212.51 | -1,094.0 | 1,224.0 |
| GCc1 | 60 | 31 | -0.0457 | 0 | 0.7951 | -1.0000 | 1.0000 | 295 | 22.21 | 16.00 | 255.19 | -1,224.0 | 1,343.0 |
| GCc1 | 120 | 44 | -0.1231 | 0 | 0.7609 | -1.0000 | 1.0000 | 295 | 68.87 | 33.00 | 325.72 | -1,068.0 | 1,727.0 |
| GCc1 | 300 | 101 | -0.1223 | 0 | 0.7797 | -1.0000 | 1.0000 | 295 | 179.89 | 70.00 | 541.19 | -836.00 | 2,904.0 |
| GCc1 | 600 | 134 | -0.1000 | 0 | 0.7143 | -1.0000 | 1.0000 | 295 | 193.20 | 64.00 | 714.03 | -2,148.0 | 3,574.0 |
| ESc1 | 5 | 314 | 0.0187 | 0.0401 | 0.3865 | -0.9574 | 1.0000 | 325 | 27.29 | 19.00 | 291.19 | -2,643.0 | 1,129.0 |
| ESc1 | 10 | 315 | 0.0139 | 0.0366 | 0.3043 | -0.9184 | 0.8769 | 325 | 46.00 | 15.00 | 359.93 | -2,340.0 | 1,470.0 |
| ESc1 | 30 | 315 | 0.0135 | 0.0337 | 0.1930 | -0.5780 | 0.5614 | 325 | 178.58 | 91.00 | 691.99 | -1,760.0 | 7,354.0 |
| ESc1 | 60 | 315 | 0.0145 | 0.0287 | 0.1485 | -0.4517 | 0.5234 | 325 | 280.69 | 111.00 | 1,224.6 | -3,108.0 | 9,153.0 |
| ESc1 | 120 | 315 | 0.0226 | 0.0340 | 0.1041 | -0.3548 | 0.3128 | 325 | 416.80 | 208.00 | 1,598.9 | -3,355.0 | 10,680.0 |
| ESc1 | 300 | 315 | 0.0215 | 0.0325 | 0.0873 | -0.2958 | 0.3043 | 325 | 905.79 | 261.00 | 2,829.5 | -5,087.0 | 20,110.0 |
| ESc1 | 600 | 315 | 0.0133 | 0.0207 | 0.0654 | -0.2415 | 0.1682 | 325 | 908.91 | 463.00 | 4,124.7 | -6,822.0 | 29,971.0 |
| BTCc1 | 5 | 194 | 0.0177 | 0 | 0.7494 | -1.0000 | 1.0000 | 320 | 2.3656 | 0 | 54.65 | -272.00 | 490.00 |
| BTCc1 | 10 | 234 | 0.0089 | 0.0358 | 0.6592 | -1.0000 | 1.0000 | 320 | 3.6812 | 0 | 76.26 | -410.00 | 389.00 |
| BTCc1 | 30 | 277 | 0.0260 | 0 | 0.5455 | -1.0000 | 1.0000 | 320 | -1.0344 | 0 | 150.65 | -760.00 | 748.00 |
| BTCc1 | 60 | 295 | -0.0143 | -0.0137 | 0.4585 | -1.0000 | 1.0000 | 320 | -11.70 | 0 | 216.65 | -1,003.0 | 991.00 |
| BTCc1 | 120 | 308 | 0.0256 | 0 | 0.3944 | -1.0000 | 1.0000 | 320 | -12.26 | 6.0000 | 291.27 | -1,155.0 | 1,101.0 |
| BTCc1 | 300 | 314 | 0.0409 | 0.0368 | 0.3524 | -1.0000 | 1.0000 | 320 | -15.78 | 18.00 | 459.87 | -1,928.0 | 1,387.0 |
| BTCc1 | 600 | 319 | 0.0282 | 0.0258 | 0.2785 | -1.0000 | 1.0000 | 320 | 25.44 | 42.00 | 600.89 | -2,457.0 | 1,644.0 |

**Table 5. Volatility Ratios by Ticker and Return Horizon**

For each FOMC announcement, the volatility ratio is post-event volatility divided by pre-event volatility. Pre-event volatility is calculated over the 30 minutes before the announcement; the 15-minute post-event window begins after the return horizon $h$ and order-flow window $w$ used to estimate $\mu$. These event-specific ratios are then averaged across valid FOMC announcements separately for each ticker–$h$–$w$ specification. The table reports the distribution of these cross-event averages across order-flow windows for each ticker and return horizon. $N$ is therefore the number of valid order-flow-window averages for that ticker and return horizon (at most the number of order-flow windows), not the number of FOMC announcements used inside each average.

| Ticker | $h$ | $N$ | Mean | Median | SD | Min | Max |
|---|---|---|---|---|---|---|---|
| CLc1 | 1 | 7 | 1.6140 | 1.6364 | 0.0857 | 1.4828 | 1.7024 |
| CLc1 | 5 | 7 | 1.6110 | 1.6337 | 0.0846 | 1.4819 | 1.6981 |
| CLc1 | 10 | 7 | 1.6082 | 1.6296 | 0.0837 | 1.4798 | 1.6945 |
| CLc1 | 30 | 7 | 1.5905 | 1.6185 | 0.0719 | 1.4717 | 1.6607 |
| CLc1 | 60 | 7 | 1.5751 | 1.5916 | 0.0602 | 1.4717 | 1.6337 |
| GCc1 | 1 | 7 | 11.65 | 12.86 | 2.2691 | 7.2706 | 13.47 |
| GCc1 | 5 | 7 | 11.74 | 12.86 | 2.2906 | 7.2712 | 13.46 |
| GCc1 | 10 | 7 | 11.45 | 12.84 | 3.0169 | 6.9071 | 13.99 |
| GCc1 | 30 | 7 | 11.20 | 12.86 | 2.9194 | 6.8821 | 13.46 |
| GCc1 | 60 | 7 | 10.45 | 12.86 | 3.9617 | 3.0926 | 13.19 |
| ESc1 | 1 | 7 | 2.3059 | 2.3588 | 0.3732 | 1.7382 | 2.7515 |
| ESc1 | 5 | 7 | 2.2786 | 2.3460 | 0.3511 | 1.7322 | 2.6773 |
| ESc1 | 10 | 7 | 2.2550 | 2.3303 | 0.3318 | 1.7284 | 2.6161 |
| ESc1 | 30 | 7 | 2.1873 | 2.2752 | 0.2829 | 1.7186 | 2.4639 |
| ESc1 | 60 | 7 | 2.1190 | 2.2012 | 0.2434 | 1.7053 | 2.3460 |
| BTCc1 | 1 | 7 | 2.0743 | 2.1339 | 0.2713 | 1.6602 | 2.3616 |
| BTCc1 | 5 | 7 | 2.0623 | 2.1228 | 0.2628 | 1.6573 | 2.3492 |
| BTCc1 | 10 | 7 | 2.0482 | 2.1120 | 0.2552 | 1.6422 | 2.3091 |
| BTCc1 | 30 | 7 | 1.9980 | 2.0688 | 0.2234 | 1.6275 | 2.2148 |
| BTCc1 | 60 | 7 | 1.9421 | 2.0114 | 0.1997 | 1.5956 | 2.1228 |

**Table 6. Measure of feedback trading susceptibility in BTCc1 Futures: Depth-Scaled Quote OFI**
This table reports estimates of $\mu$ from regressions of subsequent order flow on the post-FOMC log return, expressed in percentage points. Rows indicate the return horizon $h$ and columns indicate the subsequent order-flow window $w$. Quote OFI is constructed following Cont et al. (2014) as the cumulative sum of signed best-bid and best-ask update contributions. In this paper, we use an extended, depth-scaled version of this measure: raw quote OFI is divided by average displayed depth per side over the same order-flow window. This scaling preserves the direction of quote pressure while making its magnitude more comparable across markets with different quoted depth. For this table: $N = 64$ FOMC events in every horizon-window specification. HC1 heteroskedasticity-robust standard errors are in parentheses. $^{***}$, $^{**}$, and $^{*}$ denote significance at the 1%, 5%, and 10% levels, respectively.

| Return horizon | $w_5$ | $w_{10}$ | $w_{30}$ | $w_{60}$ | $w_{120}$ | $w_{300}$ | $w_{600}$ |
|---|---|---|---|---|---|---|---|
| $h_1$ | 2.251 | -345.962 | -365.834 | -624.713 | -574.228 | -1426.433 | 241.711 |
| | (76.977) | (535.513) | (874.430) | (1079.942) | (1128.446) | (943.595) | (1102.111) |
| $h_5$ | -6.848 | -37.934 | 33.550 | 119.744 | 186.960 | 276.744 | 308.251 |
| | (19.315) | (33.192) | (58.143) | (118.976) | (144.195) | (212.850) | (239.523) |
| $h_{10}$ | -33.746 | -20.006 | 95.708* | 170.778* | 226.792** | 246.765 | 149.603 |
| | (29.385) | (45.319) | (54.645) | (97.786) | (103.894) | (176.732) | (192.271) |
| $h_{30}$ | -4.833 | 2.417 | -2.215 | 34.442 | 55.303 | 63.989 | 41.657 |
| | (6.662) | (13.292) | (22.525) | (41.851) | (54.212) | (88.835) | (86.296) |
| $h_{60}$ | 11.564** | 19.016** | 36.224** | 36.284* | 53.564* | 109.485** | 178.660*** |
| | (5.414) | (7.437) | (14.894) | (19.375) | (28.445) | (48.905) | (57.438) |

**Table 7. Measure of feedback trading susceptibility in CLc1 Futures: Depth-Scaled Quote OFI**
This table reports estimates of $\mu$ from regressions of subsequent order flow on the post-FOMC log return, expressed in percentage points. Rows indicate the return horizon $h$ and columns indicate the subsequent order-flow window $w$. Quote OFI is constructed following Cont et al. (2014) as the cumulative sum of signed best-bid and best-ask update contributions. In this paper, we use an extended, depth-scaled version of this measure: raw quote OFI is divided by average displayed depth per side over the same order-flow window. This scaling preserves the direction of quote pressure while making its magnitude more comparable across markets with different quoted depth. For this table: $N = 64$ FOMC events in every horizon-window specification. HC1 heteroskedasticity-robust standard errors are in parentheses. $^{***}$, $^{**}$, and $^{*}$ denote significance at the 1%, 5%, and 10% levels, respectively.

| Return horizon | $w_5$ | $w_{10}$ | $w_{30}$ | $w_{60}$ | $w_{120}$ | $w_{300}$ | $w_{600}$ |
|---|---|---|---|---|---|---|---|
| $h_1$ | 6.531 | 39.053 | 170.195 | 4.891 | -510.723 | -595.656 | -646.618 |
| | (87.927) | (97.775) | (141.842) | (200.758) | (313.592) | (429.011) | (549.571) |
| $h_5$ | 11.028 | 32.042 | 17.413 | 57.484 | 88.701 | 133.499 | 95.396 |
| | (18.699) | (22.785) | (32.647) | (42.002) | (61.873) | (158.400) | (212.502) |
| $h_{10}$ | 0.623 | 17.009 | 25.009 | 54.002* | 76.853** | 108.535 | 114.242 |
| | (20.766) | (22.184) | (24.584) | (32.351) | (36.028) | (122.360) | (153.128) |
| $h_{30}$ | -5.294 | 12.239* | 20.858 | -37.699 | -64.967 | -66.921 | -78.263 |
| | (6.711) | (7.313) | (14.605) | (22.924) | (65.505) | (84.062) | (99.293) |
| $h_{60}$ | -6.983 | 4.424 | -29.028* | -35.866 | -49.249 | -96.846 | -107.813 |
| | (7.381) | (9.255) | (17.055) | (44.943) | (51.068) | (75.152) | (77.761) |

**Table 8. Measure of feedback trading susceptibility in ESc1 Futures: Depth-Scaled Quote OFI**
This table reports estimates of $\mu$ from regressions of subsequent order flow on the post-FOMC log return, expressed in percentage points. Rows indicate the return horizon $h$ and columns indicate the subsequent order-flow window $w$. Quote OFI is constructed following Cont et al. (2014) as the cumulative sum of signed best-bid and best-ask update contributions. In this paper, we use an extended, depth-scaled version of this measure: raw quote OFI is divided by average displayed depth per side over the same order-flow window. This scaling preserves the direction of quote pressure while making its magnitude more comparable across markets with different quoted depth. For this table: $N = 62$ FOMC events in every horizon-window specification. HC1 heteroskedasticity-robust standard errors are in parentheses. $^{***}$, $^{**}$, and $^{*}$ denote significance at the 1%, 5%, and 10% levels, respectively.

| Return horizon | $w_5$ | $w_{10}$ | $w_{30}$ | $w_{60}$ | $w_{120}$ | $w_{300}$ | $w_{600}$ |
|---|---|---|---|---|---|---|---|
| $h_1$ | -10.159 | 162.088 | 36.218 | -231.648 | 67.812 | -68.766 | 343.485 |
| | (90.252) | (144.173) | (161.612) | (248.221) | (312.496) | (500.096) | (670.003) |
| $h_5$ | 6.352 | -47.640 | 59.696 | 121.668 | 125.732 | 415.719 | 674.946 |
| | (40.404) | (43.340) | (57.623) | (99.659) | (162.197) | (331.985) | (562.794) |
| $h_{10}$ | -56.143** | -73.617* | -19.081 | 103.924 | 64.329 | 421.819 | 664.679 |
| | (22.390) | (40.699) | (46.168) | (88.102) | (144.954) | (323.471) | (555.014) |
| $h_{30}$ | -35.935 | -7.621 | 80.851 | 81.058 | 93.301 | 487.078** | 692.833 |
| | (30.259) | (15.551) | (55.439) | (67.332) | (128.082) | (213.052) | (430.388) |
| $h_{60}$ | 11.406 | 13.609 | -4.090 | -37.275 | 111.606 | 250.068** | 435.505 |
| | (8.504) | (10.350) | (25.278) | (58.681) | (118.408) | (127.099) | (275.211) |

**Table 9. Measure of feedback trading susceptibility in GCc1 Futures: Depth-Scaled Quote OFI**
This table reports estimates of $\mu$ from regressions of subsequent order flow on the post-FOMC log return, expressed in percentage points. Rows indicate the return horizon $h$ and columns indicate the subsequent order-flow window $w$. Quote OFI is constructed following Cont et al. (2014) as the cumulative sum of signed best-bid and best-ask update contributions. In this paper, we use an extended, depth-scaled version of this measure: raw quote OFI is divided by average displayed depth per side over the same order-flow window. This scaling preserves the direction of quote pressure while making its magnitude more comparable across markets with different quoted depth. For this table: $N = 21$ FOMC events in every horizon-window specification. HC1 heteroskedasticity-robust standard errors are in parentheses. $^{***}$, $^{**}$, and $^{*}$ denote significance at the 1%, 5%, and 10% levels, respectively.

| Return horizon | $w_5$ | $w_{10}$ | $w_{30}$ | $w_{60}$ | $w_{120}$ | $w_{300}$ | $w_{600}$ |
|---|---|---|---|---|---|---|---|
| $h_1$ | | | | | | | |
| $h_5$ | -168.523*** | -456.130*** | -897.069*** | -1198.922*** | -1167.256*** | -893.562*** | -1019.768*** |
| | (26.677) | (26.487) | (39.078) | (53.806) | (118.718) | (234.685) | (75.282) |
| $h_{10}$ | -64.774 | -151.168 | -162.412 | -308.732 | -230.559 | -59.403 | 306.568 |
| | (91.559) | (214.113) | (244.406) | (345.712) | (348.548) | (387.378) | (502.688) |
| $h_{30}$ | 4.012 | 0.767 | 24.129 | 16.977 | 58.513 | 124.979 | 126.319 |
| | (13.618) | (37.389) | (25.080) | (59.768) | (69.738) | (116.371) | (210.882) |
| $h_{60}$ | 46.789*** | 27.199 | -37.480 | -67.918** | 59.966 | 88.832 | 265.429 |
| | (11.946) | (25.246) | (88.312) | (31.862) | (59.810) | (137.070) | (248.869) |

**Table 10. Measure of feedback trading susceptibility in BTCc1 Futures: Robustness OFI Measures**

This table reports robustness estimates of $\mu$ from regressions of subsequent order flow on the post-FOMC log return, expressed in percentage points. Rows indicate the return horizon $h$ and columns indicate the subsequent order-flow window $w$. Panel A uses executed order-flow imbalance, defined as cumulative buyer-initiated volume minus cumulative seller-initiated volume, divided by total executed volume over the same window. Panel B uses unscaled quote OFI, constructed following Cont et al. (2014) as the cumulative sum of signed best-bid and best-ask update contributions. In the main analysis of this paper, we extend the Cont et al. (2014) measure by dividing raw quote OFI by average displayed depth per side over the same order-flow window; Panel B deliberately retains the unscaled version as a robustness check. The number of usable FOMC events, $N$, is reported beneath every return-horizon row in both panels to keep the side-by-side presentation consistent. HC1 heteroskedasticity-robust standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

*Panel A: Executed Order-Flow Imbalance*

| $h$ | $w_5$ | $w_{10}$ | $w_{30}$ | $w_{60}$ | $w_{120}$ | $w_{300}$ | $w_{600}$ |
|---|---|---|---|---|---|---|---|
| $h_1$ | 3.550 | 3.606* | 1.963 | 0.267 | 0.989 | 0.276 | 0.055 |
| | (2.612) | (1.863) | (1.461) | (1.183) | (0.863) | (0.575) | (0.458) |
| $N$ | 36 | 47 | 54 | 58 | 62 | 63 | 64 |
| $h_5$ | 0.197 | 0.275 | 0.293** | 0.284 | 0.366*** | 0.223*** | 0.132** |
| | (0.318) | (0.237) | (0.144) | (0.177) | (0.128) | (0.073) | (0.065) |
| $N$ | 37 | 48 | 54 | 58 | 62 | 63 | 64 |
| $h_{10}$ | 0.137 | 0.150 | 0.294** | 0.180 | 0.195** | 0.154*** | 0.072* |
| | (0.282) | (0.203) | (0.131) | (0.125) | (0.081) | (0.057) | (0.043) |
| $N$ | 38 | 42 | 54 | 58 | 62 | 63 | 64 |
| $h_{30}$ | 0.239 | 0.075 | 0.030 | -0.010 | -0.005 | -0.016 | -0.022 |
| | (0.176) | (0.104) | (0.070) | (0.044) | (0.043) | (0.030) | (0.021) |
| $N$ | 44 | 50 | 56 | 60 | 61 | 62 | 63 |
| $h_{60}$ | 0.016 | 0.032 | 0.068 | 0.029 | 0.007 | 0.014 | 0.006 |
| | (0.117) | (0.070) | (0.049) | (0.051) | (0.029) | (0.019) | (0.014) |
| $N$ | 39 | 47 | 59 | 61 | 61 | 63 | 64 |

*Panel B: Unscaled Quote OFI*

| $h$ | $w_5$ | $w_{10}$ | $w_{30}$ | $w_{60}$ | $w_{120}$ | $w_{300}$ | $w_{600}$ |
|---|---|---|---|---|---|---|---|
| $h_1$ | 45.891 | -316.384 | -500.853 | -1004.592 | -884.818 | -2266.344 | 761.594 |
| | (118.179) | (633.269) | (1228.504) | (1651.643) | (1877.298) | (1741.471) | (2063.749) |
| $N$ | 64 | 64 | 64 | 64 | 64 | 64 | 64 |
| $h_5$ | -21.050 | -76.105 | 42.711 | 143.384 | 268.196 | 474.044 | 593.970 |
| | (24.515) | (53.232) | (90.468) | (185.561) | (240.648) | (360.273) | (443.842) |
| $N$ | 64 | 64 | 64 | 64 | 64 | 64 | 64 |
| $h_{10}$ | -51.821 | -28.680 | 157.103* | 266.587* | 358.538** | 403.050 | 280.025 |
| | (39.308) | (57.610) | (92.466) | (159.183) | (169.098) | (297.178) | (350.991) |
| $N$ | 64 | 64 | 64 | 64 | 64 | 64 | 64 |
| $h_{30}$ | -4.562 | 1.542 | -23.075 | 54.674 | 87.440 | 100.369 | 73.486 |
| | (9.596) | (23.772) | (41.205) | (68.313) | (86.671) | (150.564) | (154.010) |
| $N$ | 64 | 64 | 64 | 64 | 64 | 64 | 64 |
| $h_{60}$ | 19.773** | 35.600*** | 71.564*** | 63.037** | 96.501** | 204.269** | 353.410*** |
| | (9.421) | (13.335) | (25.505) | (31.865) | (48.162) | (91.653) | (114.170) |
| $N$ | 64 | 64 | 64 | 64 | 64 | 64 | 64 |

**Table 11. Measure of feedback trading susceptibility in CLc1 Futures: Robustness OFI Measures**

This table reports robustness estimates of $\mu$ from regressions of subsequent order flow on the post-FOMC log return, expressed in percentage points. Rows indicate the return horizon $h$ and columns indicate the subsequent order-flow window $w$. Panel A uses executed order-flow imbalance, defined as cumulative buyer-initiated volume minus cumulative seller-initiated volume, divided by total executed volume over the same window. Panel B uses unscaled quote OFI, constructed following Cont et al. (2014) as the cumulative sum of signed best-bid and best-ask update contributions. In the main analysis of this paper, we extend the Cont et al. (2014) measure by dividing raw quote OFI by average displayed depth per side over the same order-flow window; Panel B deliberately retains the unscaled version as a robustness check. The number of usable FOMC events, $N$, is reported beneath every return-horizon row in both panels to keep the side-by-side presentation consistent. HC1 heteroskedasticity-robust standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

*Panel A: Executed Order-Flow Imbalance*

| $h$ | $w_5$ | $w_{10}$ | $w_{30}$ | $w_{60}$ | $w_{120}$ | $w_{300}$ | $w_{600}$ |
|---|---|---|---|---|---|---|---|
| $h_1$ | -3.174 | -2.165 | -0.936 | -0.902 | -1.077 | -0.801 | -0.288 |
| | (2.872) | (2.129) | (1.396) | (1.109) | (0.729) | (0.663) | (0.629) |
| $N$ | 61 | 64 | 64 | 64 | 64 | 64 | 64 |
| $h_5$ | 0.159 | 0.199 | -0.123 | 0.295** | 0.118 | -0.029 | 0.068 |
| | (0.588) | (0.523) | (0.195) | (0.135) | (0.093) | (0.112) | (0.086) |
| $N$ | 62 | 64 | 64 | 64 | 64 | 64 | 64 |
| $h_{10}$ | -0.143 | -0.312 | -0.161 | 0.085 | 0.007 | -0.056 | 0.008 |
| | (0.315) | (0.351) | (0.221) | (0.083) | (0.084) | (0.090) | (0.053) |
| $N$ | 60 | 62 | 64 | 64 | 64 | 64 | 64 |
| $h_{30}$ | -0.172 | -0.019 | 0.126 | 0.047 | 0.035 | 0.022 | 0.072* |
| | (0.282) | (0.253) | (0.117) | (0.121) | (0.062) | (0.054) | (0.041) |
| $N$ | 64 | 64 | 64 | 64 | 64 | 64 | 64 |
| $h_{60}$ | 0.023 | 0.098 | 0.006 | 0.021 | -0.032 | -0.023 | 0.008 |
| | (0.159) | (0.136) | (0.100) | (0.073) | (0.059) | (0.041) | (0.041) |
| $N$ | 63 | 63 | 64 | 64 | 64 | 64 | 64 |

*Panel B: Unscaled Quote OFI*

| $h$ | $w_5$ | $w_{10}$ | $w_{30}$ | $w_{60}$ | $w_{120}$ | $w_{300}$ | $w_{600}$ |
|---|---|---|---|---|---|---|---|
| $h_1$ | 254.779 | 219.172 | 1244.222 | 423.372 | -2351.595 | -5708.553** | -8962.841*** |
| | (932.024) | (785.244) | (1102.831) | (1505.471) | (2432.394) | (2768.274) | (3251.873) |
| $N$ | 64 | 64 | 64 | 64 | 64 | 64 | 64 |
| $h_5$ | 94.629 | 174.513* | 303.658* | 392.595 | 850.714** | 442.190 | -323.331 |
| | (86.828) | (98.223) | (173.402) | (282.072) | (347.395) | (902.172) | (1221.188) |
| $N$ | 64 | 64 | 64 | 64 | 64 | 64 | 64 |
| $h_{10}$ | -20.573 | 89.097 | 292.998 | 324.729 | 590.777** | 592.125 | 407.415 |
| | (113.001) | (129.149) | (204.777) | (208.170) | (244.217) | (803.018) | (1015.554) |
| $N$ | 64 | 64 | 64 | 64 | 64 | 64 | 64 |
| $h_{30}$ | 12.948 | 164.268*** | 109.985 | -222.080 | -389.196 | -581.377 | -877.547 |
| | (48.206) | (62.559) | (89.732) | (154.775) | (399.523) | (472.901) | (561.642) |
| $N$ | 64 | 64 | 64 | 64 | 64 | 64 | 64 |
| $h_{60}$ | -58.124* | 21.615 | -174.973* | -210.177 | -375.463 | -658.058 | -819.708* |
| | (31.802) | (43.856) | (104.818) | (262.092) | (308.901) | (447.522) | (476.831) |
| $N$ | 64 | 64 | 64 | 64 | 64 | 64 | 64 |

**Table 12. Measure of feedback trading susceptibility in ESc1 Futures: Robustness OFI Measures**

This table reports robustness estimates of $\mu$ from regressions of subsequent order flow on the post-FOMC log return, expressed in percentage points. Rows indicate the return horizon $h$ and columns indicate the subsequent order-flow window $w$. Panel A uses executed order-flow imbalance, defined as cumulative buyer-initiated volume minus cumulative seller-initiated volume, divided by total executed volume over the same window. Panel B uses unscaled quote OFI, constructed following Cont et al. (2014) as the cumulative sum of signed best-bid and best-ask update contributions. In the main analysis of this paper, we extend the Cont et al. (2014) measure by dividing raw quote OFI by average displayed depth per side over the same order-flow window; Panel B deliberately retains the unscaled version as a robustness check. The number of usable FOMC events, $N$, is reported beneath every return-horizon row in both panels to keep the side-by-side presentation consistent. HC1 heteroskedasticity-robust standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

*Panel A: Executed Order-Flow Imbalance*

| $h$ | $w_5$ | $w_{10}$ | $w_{30}$ | $w_{60}$ | $w_{120}$ | $w_{300}$ | $w_{600}$ |
|---|---|---|---|---|---|---|---|
| $h_1$ | 1.157 | 0.794 | 0.370 | 0.205 | 0.308 | 0.227 | 0.174 |
| | (1.077) | (1.037) | (0.606) | (0.530) | (0.305) | (0.177) | (0.143) |
| $N$ | 62 | 63 | 63 | 63 | 63 | 63 | 63 |
| $h_5$ | 0.671*** | 0.556*** | 0.486*** | 0.406*** | 0.212*** | 0.163*** | 0.082* |
| | (0.200) | (0.169) | (0.097) | (0.084) | (0.078) | (0.055) | (0.045) |
| $N$ | 63 | 63 | 63 | 63 | 63 | 63 | 63 |
| $h_{10}$ | 0.319* | 0.232** | 0.235*** | 0.255*** | 0.143*** | 0.092** | 0.034 |
| | (0.168) | (0.115) | (0.072) | (0.058) | (0.047) | (0.039) | (0.031) |
| $N$ | 63 | 63 | 63 | 63 | 63 | 63 | 63 |
| $h_{30}$ | 0.225 | 0.143 | 0.175*** | 0.133** | 0.058 | 0.036 | 0.012 |
| | (0.178) | (0.107) | (0.067) | (0.067) | (0.040) | (0.031) | (0.023) |
| $N$ | 63 | 63 | 63 | 63 | 63 | 63 | 63 |
| $h_{60}$ | 0.191** | 0.141*** | 0.062 | 0.024 | 0.015 | 0.016 | 0.003 |
| | (0.089) | (0.053) | (0.058) | (0.033) | (0.031) | (0.025) | (0.019) |
| $N$ | 63 | 63 | 63 | 63 | 63 | 63 | 63 |

*Panel B: Unscaled Quote OFI*

| $h$ | $w_5$ | $w_{10}$ | $w_{30}$ | $w_{60}$ | $w_{120}$ | $w_{300}$ | $w_{600}$ |
|---|---|---|---|---|---|---|---|
| $h_1$ | -1358.866 | 141.523 | -36.966 | 107.534 | 101.690 | 196.795 | 1073.703 |
| | (949.494) | (1280.114) | (911.008) | (1905.175) | (2218.899) | (4557.823) | (5422.148) |
| $N$ | 63 | 63 | 63 | 63 | 63 | 63 | 63 |
| $h_5$ | 1.452 | -259.793 | 193.029 | 2076.579 | 1822.778 | 5519.382 | 6025.957 |
| | (165.713) | (255.718) | (312.323) | (1649.076) | (2057.290) | (4158.569) | (6391.882) |
| $N$ | 63 | 63 | 63 | 63 | 63 | 63 | 63 |
| $h_{10}$ | -235.240 | -274.816 | -258.069 | 2170.657 | 1731.586 | 5949.330 | 7688.695 |
| | (154.104) | (202.744) | (282.452) | (1830.487) | (2036.401) | (3941.534) | (6191.216) |
| $N$ | 63 | 63 | 63 | 63 | 63 | 63 | 63 |
| $h_{30}$ | -60.706 | -48.626 | 2050.747* | 2352.933 | 2124.018 | 6160.816** | 8044.142* |
| | (113.146) | (75.522) | (1123.691) | (1507.321) | (1787.217) | (2675.629) | (4816.505) |
| $N$ | 63 | 63 | 63 | 63 | 63 | 63 | 63 |
| $h_{60}$ | 129.718** | 282.302 | 209.228 | -113.715 | 1851.280 | 2782.058** | 4516.864 |
| | (52.665) | (181.413) | (318.783) | (546.760) | (1538.840) | (1323.633) | (2831.131) |
| $N$ | 63 | 63 | 63 | 63 | 63 | 63 | 63 |

**Table 13. Measure of feedback trading susceptibility in GCc1 Futures: Robustness OFI Measures**

This table reports robustness estimates of $\mu$ from regressions of subsequent order flow on the post-FOMC log return, expressed in percentage points. Rows indicate the return horizon $h$ and columns indicate the subsequent order-flow window $w$. Panel A uses executed order-flow imbalance, defined as cumulative buyer-initiated volume minus cumulative seller-initiated volume, divided by total executed volume over the same window. Panel B uses unscaled quote OFI, constructed following Cont et al. (2014) as the cumulative sum of signed best-bid and best-ask update contributions. In the main analysis of this paper, we extend the Cont et al. (2014) measure by dividing raw quote OFI by average displayed depth per side over the same order-flow window; Panel B deliberately retains the unscaled version as a robustness check. The number of usable FOMC events, $N$, is reported beneath every return-horizon row in both panels to keep the side-by-side presentation consistent. HC1 heteroskedasticity-robust standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

*Panel A: Executed Order-Flow Imbalance*

| $h$ | $w_5$ | $w_{10}$ | $w_{30}$ | $w_{60}$ | $w_{120}$ | $w_{300}$ | $w_{600}$ |
|---|---|---|---|---|---|---|---|
| $h_1$ | | | | | | | |
| $N$ | | | | | | | |
| $h_5$ | | | | | | -1.526 | -0.496 |
| | | | | | | (1.477) | (1.173) |
| $N$ | | | | | | 14 | 15 |
| $h_{10}$ | | | | | | -2.165** | -1.723 |
| | | | | | | (0.940) | (1.136) |
| $N$ | | | | | | 13 | 14 |
| $h_{30}$ | | | | | | 0.066 | -0.487 |
| | | | | | | (0.769) | (0.774) |
| $N$ | | | | | | 11 | 12 |
| $h_{60}$ | | | | | | 0.415 | -0.245 |
| | | | | | | (1.368) | (1.503) |
| $N$ | | | | | | 12 | 13 |

*Panel B: Unscaled Quote OFI*

| $h$ | $w_5$ | $w_{10}$ | $w_{30}$ | $w_{60}$ | $w_{120}$ | $w_{300}$ | $w_{600}$ |
|---|---|---|---|---|---|---|---|
| $h_1$ | | | | | | | |
| $N$ | | | | | | | |
| $h_5$ | -867.955*** | -2495.006*** | -5551.980*** | -6009.990*** | -5123.652*** | -6348.330* | -6062.795** |
| | (81.754) | (150.664) | (1346.250) | (1395.963) | (967.800) | (3834.875) | (2454.618) |
| $N$ | 21 | 21 | 21 | 21 | 21 | 21 | 21 |
| $h_{10}$ | -506.515 | -914.607 | -1211.547 | -1590.491 | -1224.081 | -1523.646 | 65.385 |
| | (474.351) | (876.087) | (1129.739) | (1392.028) | (1283.635) | (2032.435) | (2238.017) |
| $N$ | 21 | 21 | 21 | 21 | 21 | 21 | 21 |
| $h_{30}$ | 170.325 | 21.971 | 175.005 | 103.452 | 456.517 | 1008.051 | 939.705 |
| | (191.390) | (150.991) | (158.057) | (177.735) | (394.158) | (832.971) | (938.529) |
| $N$ | 21 | 21 | 21 | 21 | 21 | 21 | 21 |
| $h_{60}$ | 332.515*** | 140.981* | -206.121 | -368.979** | 621.551 | 1240.964 | 1738.507** |
| | (74.468) | (81.597) | (285.249) | (160.111) | (392.581) | (940.074) | (761.820) |
| $N$ | 21 | 21 | 21 | 21 | 21 | 21 | 21 |

**Table 14. Financial Instability and Standardized Mu**

The dependent variable is the post-event to pre-event volatility ratio. Within each order-flow group, column (1) reports pooled OLS, column (2) includes ticker fixed effects, and column (3) includes ticker, return-horizon, and order-flow-window fixed effects. The three groups use depth-scaled quote OFI, executed OFI, and unscaled quote OFI, respectively, to estimate $\mu$. HC1 heteroskedasticity-robust standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

| | Depth-Scaled Quote OFI | | | Executed OFI | | | Unscaled Quote OFI | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
| Standardized $\widehat{\mu}$ | -2.304*** | -0.634*** | -0.281* | 0.568 | 0.753* | 0.526 | -1.658*** | -0.453** | -0.285* |
| | (0.658) | (0.202) | (0.164) | (1.113) | (0.395) | (0.465) | (0.639) | (0.195) | (0.146) |
| Constant | 2.556 | 2.053 | 2.279 | 1.790 | 1.930 | 2.208 | 2.551 | 2.035 | 2.271 |
| | (0.158) | (0.056) | (0.070) | (0.160) | (0.063) | (0.062) | (0.155) | (0.057) | (0.069) |
| Ticker FE | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| Return-horizon FE | No | No | Yes | No | No | Yes | No | No | Yes |
| Order-flow-window FE | No | No | Yes | No | No | Yes | No | No | Yes |
| Observations | 133 | 133 | 133 | 113 | 113 | 113 | 133 | 133 | 133 |
| $R^2$ | 0.091 | 0.943 | 0.967 | 0.006 | 0.909 | 0.944 | 0.079 | 0.942 | 0.968 |
| Adjusted $R^2$ | 0.084 | 0.941 | 0.963 | -0.003 | 0.905 | 0.936 | 0.072 | 0.940 | 0.964 |

**Table 15. Financial Instability and Raw Mu**

The dependent variable is the post-event to pre-event volatility ratio. Within each order-flow group, column (1) reports pooled OLS, column (2) includes ticker fixed effects, and column (3) includes ticker, return-horizon, and order-flow-window fixed effects. The three groups use depth-scaled quote OFI, executed OFI, and unscaled quote OFI, respectively, to estimate $\mu$. HC1 heteroskedasticity-robust standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

| | Depth-Scaled Quote OFI | | | Executed OFI | | | Unscaled Quote OFI | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
| Raw $\widehat{\mu}$ | -0.002*** | -0.000 | -0.000 | 0.055 | 0.102** | 0.073** | -0.000** | -0.000 | -0.000 |
| | (0.001) | (0.000) | (0.000) | (0.181) | (0.049) | (0.035) | (0.000) | (0.000) | (0.000) |
| Constant | 2.439 | 1.981 | 2.271 | 1.828 | 1.947 | 2.176 | 2.492 | 1.988 | 2.276 |
| | (0.146) | (0.057) | (0.091) | (0.101) | (0.059) | (0.061) | (0.154) | (0.057) | (0.082) |
| Ticker FE | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| Return-horizon FE | No | No | Yes | No | No | Yes | No | No | Yes |
| Order-flow-window FE | No | No | Yes | No | No | Yes | No | No | Yes |
| Observations | 133 | 133 | 133 | 113 | 113 | 113 | 133 | 133 | 133 |
| $R^2$ | 0.077 | 0.940 | 0.966 | 0.002 | 0.909 | 0.945 | 0.053 | 0.940 | 0.966 |
| Adjusted $R^2$ | 0.070 | 0.938 | 0.962 | -0.007 | 0.905 | 0.937 | 0.045 | 0.938 | 0.962 |

**Table 16. Trading Behavior and Performance of LLM Agents**

The table reports results from ten independent experimental sessions with nine LLM-based agents each. RD measures the signed deviation of transaction prices from fundamental value, while RAD measures the absolute deviation. Cumulative profit is measured relative to expected initial wealth and therefore also reflects realized dividend outcomes. Excess profit relative to buy-and-hold isolates the redistribution associated with trading. Because the experiment is a closed market, aggregate excess profit from trading is zero.

| Variable | N | Mean | Median | Min | Max |
|---|---|---|---|---|---|
| *Panel A: Market-level outcomes* | | | | | |
| Trades per experimental repetition | 10 | 21.60 | 17.00 | 12.00 | 37.00 |
| Turnover (% of initial shares) | 10 | 120.00 | 94.44 | 66.67 | 205.56 |
| Hold decisions (%) | 10 | 51.78 | 56.39 | 28.89 | 63.33 |
| Buyer-initiated trades (%) | 10 | 37.06 | 36.82 | 0.00 | 93.33 |
| Aggressive buy executions | 10 | 8.60 | 9.50 | 0.00 | 18.00 |
| Aggressive sell executions | 10 | 13.00 | 13.00 | 1.00 | 21.00 |
| Relative deviation (%) | 10 | -0.315 | -0.333 | -0.625 | -0.076 |
| Relative absolute deviation (%) | 10 | 0.329 | 0.358 | 0.103 | 0.625 |
| Mean price premium (%) | 10 | -0.237 | -0.240 | -0.458 | -0.068 |
| Peak price premium (%) | 10 | 0.054 | 0.000 | 0.000 | 0.463 |
| Trades above fundamental value (%) | 10 | 1.94 | 0.00 | 0.00 | 8.33 |
| Bids above fundamental value (%) | 10 | 0.25 | 0.00 | 0.00 | 1.32 |
| Price bubble indicator | 10 | 0.00 | 0.00 | 0.00 | 0.00 |
| Realized total dividends (cents) | 10 | 392.40 | 390.00 | 288.00 | 516.00 |
| *Panel B: Agent-level outcomes* | | | | | |
| Purchases | 90 | 2.40 | 2.00 | 0.00 | 9.00 |
| Sales | 90 | 2.40 | 2.00 | 0.00 | 11.00 |
| Hold decisions (%) | 90 | 51.78 | 58.33 | 0.00 | 78.33 |
| Cumulative profit (cents) | 90 | 64.80 | 63.00 | -341.00 | 498.00 |
| Cumulative return (%) | 90 | 4.97 | 4.83 | -26.13 | 38.16 |
| Excess profit vs. buy-and-hold (cents) | 90 | 0.00 | 0.00 | -237.00 | 240.00 |
| Excess return vs. buy-and-hold (%) | 90 | 0.00 | 0.00 | -18.16 | 18.39 |

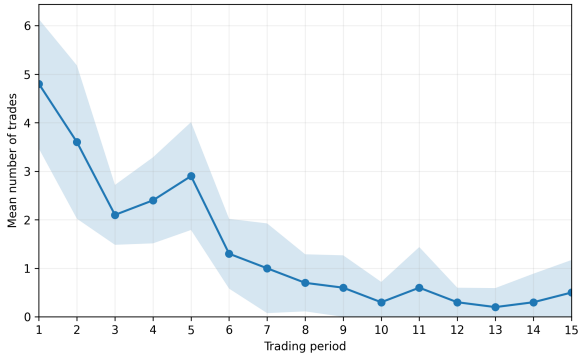
**Table 17. Trading Behavior and Performance of Reinforcement-Learning Agents in Out-of-Sample Markets**

The table reports results for ten independently trained groups of nine reinforcement-learning agents. Each group was evaluated in 100 new experimental sessions that were not used for training. To avoid treating repeated tests of the same learned strategies as independent observations, market-level values are first averaged within each trained group ($N = 10$), and trader-level values are first averaged for each trader across its 100 evaluation sessions ($N = 90$). RD measures the signed deviation of transaction prices from fundamental value, while RAD measures the absolute deviation. The price bubble indicator is the fraction of a group's evaluation sessions in which the mean transaction price remained more than 5% above fundamental value for at least two consecutive periods and was followed by a price decline of at least 5%. Cumulative profit is measured relative to expected initial wealth and therefore also reflects realized dividend outcomes. Excess profit relative to buy-and-hold isolates the redistribution associated with trading. Because the experiment is a closed market, aggregate excess profit from trading is zero.

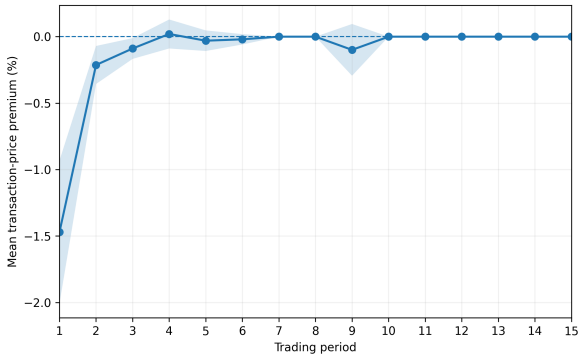
| Variable | N | Mean | Median | Min | Max |
|---|---|---|---|---|---|
| *Panel A: Market-level outcomes* | | | | | |
| Trades per experimental repetition | 10 | 8.84 | 4.77 | 1.81 | 31.44 |
| Turnover (% of initial shares) | 10 | 49.13 | 26.50 | 10.06 | 174.67 |
| Hold decisions (%) | 10 | 43.71 | 40.24 | 23.48 | 71.63 |
| Buyer-initiated trades (%) | 10 | 63.38 | 62.42 | 21.73 | 100.00 |
| Aggressive buy executions | 10 | 6.91 | 1.90 | 1.06 | 28.84 |
| Aggressive sell executions | 10 | 1.94 | 1.39 | 0.00 | 4.55 |
| Relative deviation (%) | 10 | 213.961 | 165.446 | 44.109 | 414.083 |
| Relative absolute deviation (%) | 10 | 233.099 | 200.490 | 91.655 | 414.083 |
| Mean price premium (%) | 10 | 428.046 | 320.463 | 75.293 | 1042.654 |
| Peak price premium (%) | 10 | 959.059 | 646.841 | 75.293 | 2904.704 |
| Trades above fundamental value (%) | 10 | 80.41 | 100.00 | 23.23 | 100.00 |
| Bids above fundamental value (%) | 10 | 44.11 | 42.37 | 0.00 | 85.44 |
| Price bubble indicator | 10 | 0.02 | 0.00 | 0.00 | 0.16 |
| Realized total dividends (cents) | 10 | 361.01 | 362.12 | 347.92 | 371.76 |
| *Panel B: Agent-level outcomes* | | | | | |
| Purchases | 90 | 0.98 | 0.00 | 0.00 | 13.94 |
| Sales | 90 | 0.98 | 0.00 | 0.00 | 14.28 |
| Hold decisions (%) | 90 | 41.91 | 0.00 | 0.00 | 100.00 |
| Cumulative profit (cents) | 90 | 2.02 | 2.04 | -892.75 | 1667.63 |
| Cumulative return (%) | 90 | 0.15 | 0.16 | -68.41 | 127.79 |
| Excess profit vs. buy-and-hold (cents) | 90 | 0.00 | 0.00 | -898.51 | 1663.55 |
| Excess return vs. buy-and-hold (%) | 90 | 0.00 | 0.00 | -68.85 | 127.48 |

**Figure 1: Trading behavior and performance of LLM-based agents over the experimental periods**
Panel (a) shows average trading activity, Panel (b) the average transaction-price premium relative to fundamental value, and Panel (c) the average excess profit in cents relative to buy-and-hold by initial endowment. All measures are calculated across ten independent experimental repetitions, and shaded areas represent 95% confidence intervals.

**(a)** Trading activity over time




**(b)** Transaction-price premium over time




**(c)** Trading performance over time

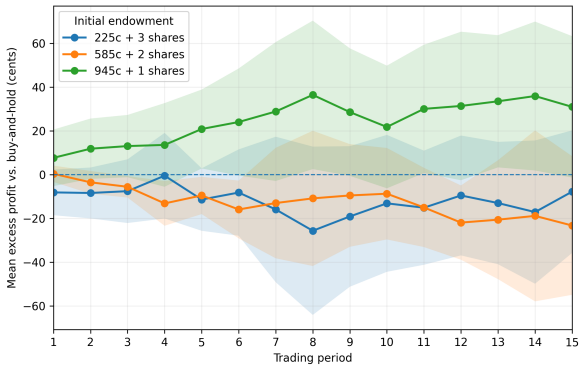
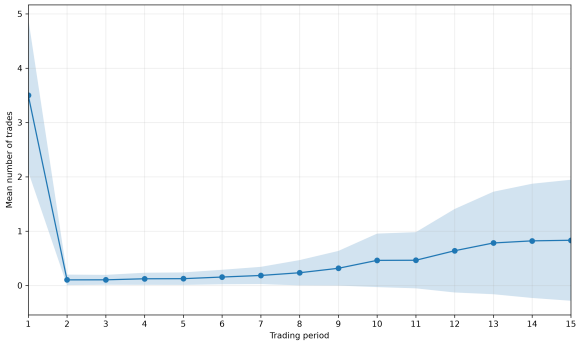
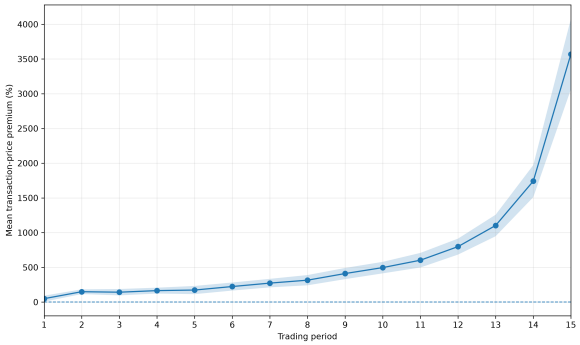


**Figure 2: Trading behavior and performance of reinforcement-learning agents over the experimental periods** Panel (a) shows average trading activity, Panel (b) the average transaction-price premium relative to fundamental value, and Panel (c) the average excess profit relative to buy-and-hold by initial endowment. Results are first averaged across 100 out-of-sample experimental repetitions for each independently trained group of nine reinforcement-learning agents and then averaged across the ten groups. Shaded areas represent 95% confidence intervals across the ten groups.

**(a)** Trading activity over time




**(b)** Transaction-price premium over time




**(c)** Trading performance over time

