How do retail investors use order flow data?

Philipp Chapkovski, Mariana Khapko, and Marius Zoican

Discussant: Kris Shen
The Ohio State University

November 20, 2025

Big Picture

What is happening in the real world?

- Retail traders now play a critical and growing role in market volume
- Demand for in-depth data products is rising across heterogeneous providers
- How retail traders value and use such data remains largely a black box
- Understanding this is essential for policy, market fairness, and data pricing

How can we study it?

- Surveys: easy, but subject to hypothetical bias
- Field data: realistic, but suffers from selection into premium feeds
- Structural models: informative, but rely on strong assumptions
- Ideal approach: a randomized experiment that varies data access cleanly

This paper: Implements exactly such a design through a randomized online experiment

This Paper in a Nutshell

Experimental Design

- Randomized online trading experiment with real-time price updates
- Participants trade with and without order flow data
- Treatments vary data quality (insider share α) and data quantity (arrival rate λ)
- Beliefs, trading behavior, and willingness to pay are measured pre- and post-trading

Main Findings

- Retail traders undervalue data relative to a Bayesian benchmark
- They capture only a small share of the theoretical payoff value
- Data access slightly improves decisions but increases over-trading
- Data mitigates the disposition effect, mainly by increasing realized losses
- Overconfidence and "premium" framing inflate willingness to pay

Comment 1: Limited Learning Time

- 30 minutes for all players to complete the game on the same day $\Longrightarrow$ willingness to pay (WTP) only before and after a very short game, 40 price updates $\times$ 8 sec each $\approx$ 320 sec, so it mainly captures initial experimentation rather than long run learning
  - Participants have limited time to develop usable heuristics that map order flow into profitable rules, which may naturally lower the observed value they extract from the data

- Real markets reveal value through subscription duration and upgrade or downgrade decisions, so the paper could emphasize that its estimates are a short run lower bound on the private value of data

- If possible, a follow-up game allowing players more time to "digest and learn" (e.g., 30 min vs 1 hour) could be informative

Comment 2: Benchmark and Control Group

- The paper benchmarks human players against a fully Bayesian trader who uses both prices and order flow, which is a demanding and potentially unrealistic standard for many retail investors
    - For players who trade similarly to the Bayesian benchmark, are we truly capturing Bayesian updating, or are we observing patterns such as trend recognition or past performance chasing?

- Alternative benchmarks could include simple heuristic rules (for example trend following based on prices or order flow), and examining how close participants come to those

- Control group: Data access in both Round #1 and Round #2 or neither

Comment 3: External Validity

- The experimental signal is a clean buy or sell order stream from bots with fixed data quality and quantity parameters, which is considerably simpler than real Level II data
  - Actual Level II feeds are denser, faster, and noisier with multiple price levels, cancellations, and cross-symbol activity
  - Players are told that the data contains *"informed traders who know the current state (they buy in good states and sell in bad states)"*, which may inflate the perceived value of data access in the game and may not fully reflect real trading environments with messier signals

- The authors might consider framing their treatment as trade history inspired by Level II, and explicitly discussing the trade off between informativeness and complexity when mapping results to real data subscriptions

Other suggestions

- Add benchmark with Player FE and Round FE for all panel regressions (e.g., Table 6) to capture the average treatment effect of data access within players and rounds
- Sequence of data rounds seems important and needs more explanation:
  - Table 7: "Data round is first" drives stronger realization of losses (disposition effect)
  - Table 8: "Data round is first" leads to less accurate beliefs
- Consider another overconfidence proxy: difference between real performance and self-assessed NASA-TLX performance (Table 9)
- A distribution figure of final WTP as a percentage of trading assets would help readers/regulators understand overall valuation of order flow data
- Other minor comments ▸ minor comments

Final Marks

- Very interesting paper with extremely careful implementation and experiment design
  - Clean causal identification
  - Novel findings: Data access mitigates the disposition effect at the cost of excessive trading
  - A paper that will benefit regulators and policy makers
- It would be helpful to:
  - Discuss (potentially improve) the limited learning time in the game and its impact on WTP estimation
  - Consider alternative benchmarks and if possible, add control groups
  - Elaborate on external validity and the informativeness–complexity tradeoff
- Good Luck!

Minor Comments ◂ Back

- Table 2 and text description do not match, "In Sessions I and III, there are few insiders ($\alpha = 0.3$)..." Need to double check the text
- Table 6 Column (7) missing obs and R^2
- Table 5 caption misused comma in equation