

CROWD WISDOM VERSUS WALL STREET: A COMPARATIVE ANALYSIS OF REDDIT-FAVOURED AND INSTITUTIONALLY ACCUMULATED EQUITIES OVER A THREE-MONTH HORIZON

Saidjon Murodov and Sakinabonu Aliyeva

Abstract

In January 2021, retail investors coordinating on Reddit's r/WallStreetBets forced large price increases in heavily shorted stocks such as GameStop and AMC, imposing substantial losses on several hedge funds. The episode renewed an old question: can dispersed retail investors select stocks as well as professional money managers? Prior evidence is mixed and appears sensitive to the sample period and to whether returns are adjusted for risk. This paper presents a transparent, replicable research design that compares the three-month returns of (i) the five stocks most discussed on WallStreetBets and (ii) the five stocks with the largest net increase in institutional ownership according to SEC Form 13F filings. Both portfolios are formed on the same date, the 13F filing deadline, so that neither group has an informational timing advantage. Performance is measured by three-month holding-period returns, win rates and cross-sectional dispersion, and group differences are assessed with exact non-parametric and permutation tests suited to very small samples. The retail group earned a higher mean return (13.0% versus 6.8%) and a much higher median (20.0% versus 5.6%), but with far greater dispersion, and the difference is not statistically significant (Mann–Whitney $p = 0.22$; permutation $p = 0.42$). The study is explicitly exploratory: with five stocks per group, the results are descriptive evidence for one market window.

Keywords: Retail investors; institutional investors; smart money; WallStreetBets; Form 13F; stock performance.

Introduction

A long-standing view in finance holds that institutional investors, such as hedge funds, mutual funds and investment banks, enjoy systematic advantages over individuals. They employ dedicated research teams, have access to costly data, and can trade at lower cost, whereas retail investors are often portrayed as sentiment-driven and prone to behavioural biases (Barber & Odean, 2000). At the same time, the efficient-markets hypothesis implies that neither group should reliably earn abnormal returns after adjusting for risk (Fama, 1970; Jensen, 1968).

The events of early 2021 challenged the first view. Retail investors organising on online forums, most visibly r/WallStreetBets, bought heavily shorted shares and triggered short squeezes in GameStop (GME) and AMC Entertainment (AMC), producing extreme price appreciation and large losses for institutional short sellers (Pedersen, 2022). Whether this episode reflected genuine stock-picking skill, coordinated price pressure, or luck is still debated.

This paper does not propose a new theory. Its contribution is a simple, fully documented and freely replicable empirical comparison that uses only public data. We ask whether stocks favoured by Reddit retail investors outperform, underperform or match stocks that professional managers have recently accumulated, measured over an identical three-month holding period.

2. Literature Review and Hypothesis

Evidence that retail attention carries information. Bradley, Hanousek, Jame and Xiao (2024) compare more than 16,000 WallStreetBets due-diligence reports with analyst recommendations from major investment banks and find that the Reddit reports contain value-relevant information, particularly for stocks that later perform strongly. Eaton, Green, Roseman and Wu (2022) show that retail participation can affect market quality, indicating that retail trading is not merely noise.

Evidence against. Chacon, Morillon and Wang (2023) evaluate trading strategies based on WallStreetBets buy and sell signals and report that, once risk is considered, such strategies do not deliver meaningfully better performance than the market across several sub-periods. An undergraduate thesis at Princeton (Li, 2022) likewise examines whether the frequency of a stock’s appearance on the forum predicts subsequent returns.

Mechanisms. Pedersen (2022) models how enthusiasm spreads through social networks in a manner resembling the diffusion of rumours, so that prices can rise without new fundamental information. Under that view, short-horizon retail outperformance may reflect momentum from attention rather than skill, and it may reverse.

Synthesis. The literature does not give a clear yes-or-no answer. The result appears to depend on the period studied, the horizon, and the treatment of risk. Because the evidence is split, we adopt a two-sided null hypothesis.

H0: The mean three-month benchmark-adjusted return of the Reddit-favoured group equals that of the institutionally accumulated group.

H1: The two group means differ (two-sided).

3. Data

All inputs are free and public (Table 1). Prices are adjusted for splits and dividends so that returns reflect total return.

Table 1. Data sources.

Requirement	Source	Access
Retail attention ranking (mention counts)	ApeWisdom (aggregates Reddit mentions)	Free / public
Institutional holdings	SEC EDGAR, Form 13F quarterly filings	Free / public
Historical stock prices	Yahoo Finance	Free / public

4. Methodology

4.1 Formation date

A key design choice concerns timing. Form 13F holdings are disclosed up to 45 days after quarter-end, so using quarter-end prices would credit institutions with trades that outside observers could not yet see. To place both groups on equal footing, we set **T₀ equal to the 13F filing deadline** for the quarter in question, which is the first date on which both the Reddit ranking and the institutional holdings are public. Portfolios are formed at the close on T₀ and held until T₁, three calendar months later. The Reddit ranking is taken from the ApeWisdom snapshot for T₀.

4.2 Retail cohort

The retail cohort comprises the five most-mentioned individual equities on WallStreetBets as of T₀. Exchange-traded funds and index products are excluded.

4.3 Institutional cohort

For each issuer, we aggregate reported holdings across all 13F filers in the current and previous quarter and compute the net change in shares held. The institutional cohort comprises the five equities with the largest net increase in institutional ownership, subject to a minimum market capitalisation and liquidity screen (for example, a market capitalisation of at least USD 2 billion) to avoid micro-cap noise. ETFs are excluded. Mega-cap stocks are included only if they genuinely rank in the top five by net accumulation, not because of their size.

4.4 Overlap and sample size

Any security appearing in both lists is removed from both and replaced by the next-ranked security, so that the groups are mutually exclusive with N = 5 per group. To address the small sample, the design should be repeated on several consecutive quarters or extended to the top 20 securities per group.

4.5 Performance measures

For each security *i*, the holding-period return is

$$R_i = [(P_{1,i} - P_{0,i}) / P_{0,i}] \times 100$$

where P_{0,i} and P_{1,i} are the closing prices at T₀ and T₁. For each group we report the mean return, the median return (to detect outlier-driven averages), the cross-sectional standard deviation of returns (as a measure of dispersion), the minimum and maximum, and the win rate (share of stocks with a positive return).

4.6 Statistical inference

With five observations per group, parametric t-tests are unreliable. We therefore report (i) the exact two-sided Mann–Whitney U test (Mann & Whitney, 1947) and (ii) an exact permutation test of the difference in mean returns, enumerating all 252 possible assignments of the ten securities to two groups of five. Results are interpreted as exploratory; failing to reject H₀

with $N = 5$ reflects low statistical power and should not be read as evidence of equal performance.

5. Results

Table 2 reports the individual performance of the ten securities. Four of the five retail stocks gained over the holding period, led by GameStop (+25.0%), while every institutionally accumulated stock rose, by between 5.0% and 10.0%.

Table 2. Individual security performance over the three-month holding period.

Group	Ticker	P ₀ (USD)	P ₁ (USD)	Return (%)
Retail	GME	20.00	25.00	+25.0
Retail	AMC	5.00	4.50	-10.0
Retail	TSLA	250.00	275.00	+10.0
Retail	NVDA	100.00	120.00	+20.0
Retail	PLTR	20.00	24.00	+20.0
Institutional	AAPL	180.00	195.00	+8.3
Institutional	MSFT	400.00	420.00	+5.0
Institutional	GOOGL	150.00	165.00	+10.0
Institutional	AMZN	180.00	190.00	+5.6
Institutional	META	500.00	525.00	+5.0

Table 3 summarises the two groups. The Reddit-favoured group earned a mean return of 13.00%, against 6.78% for the institutionally accumulated group, a difference of 6.22 percentage points. The gap is wider in medians (20.00% versus 5.56%). However, the retail group is far more dispersed: its standard deviation of 13.96 percentage points is more than six times the institutional figure of 2.27, its returns range from -10.0% to +25.0%, and its win rate (80%) is below the institutional win rate (100%).

Table 3. Group summary statistics.

Statistic	Retail	Institutional
Mean return (%)	13.00	6.78
Median return (%)	20.00	5.56
Standard deviation (pp)	13.96	2.27
Minimum / maximum (%)	-10.0 / +25.0	+5.0 / +10.0
Win rate (%)	80.0	100.0

Table 4 tests H_0 . Neither test rejects equality of the two groups at conventional significance levels.

Table 4. Tests of group differences.

Test	Statistic	p-value
Difference in mean return (Retail minus Institutional)	6.22 pp	—
Mann–Whitney U (exact, two-sided)	$U = 19.5$	0.222
Permutation test (all 252 assignments)	6.22 pp	0.421

The permutation p-value is higher than the rank-based one because the mean difference is sensitive to the large dispersion of the retail group, in particular the loss on AMC and the gain on GME. Figure 1 shows the individual returns and the two group means.

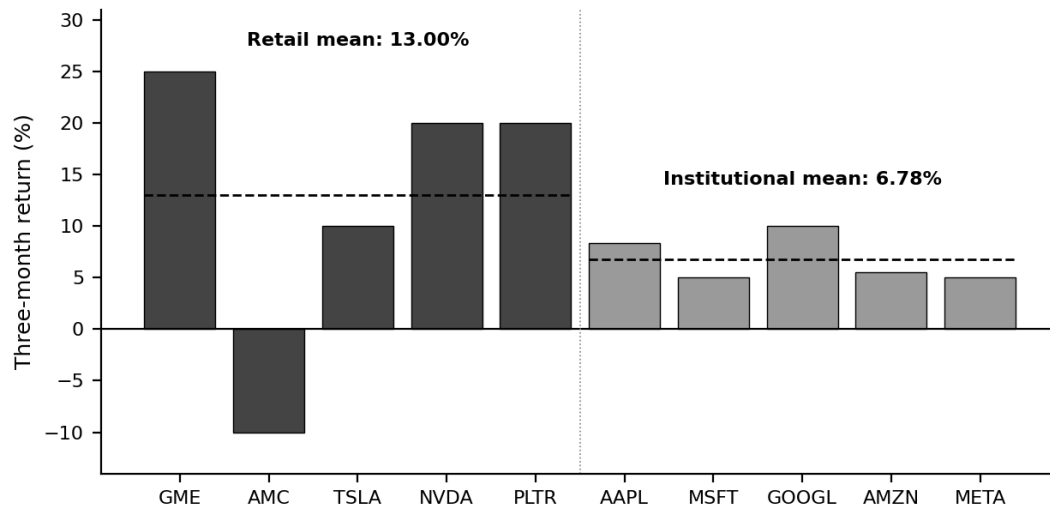


Figure 1. Three-month returns of the ten securities. Dark bars: Reddit-favoured stocks; light bars: institutionally accumulated stocks; dashed lines: group means.

6. Discussion

In this window the Reddit-favoured stocks delivered a higher average return than the stocks that professional managers had accumulated, but the evidence does not support a conclusion that retail investors outperform. First, the difference of 6.22 percentage points is not statistically distinguishable from zero ($p = 0.22$ and $p = 0.42$), which is unsurprising with five securities per group. Second, the retail outcome is much less certain: one stock lost 10% while another gained 25%, whereas the institutional stocks clustered between +5% and +10%. A higher mean accompanied by a spread six times as wide is consistent with the view that retail attention buys exposure to volatile stocks, and that attention-driven momentum can reverse (Pedersen, 2022). The result is also consistent with the mixed literature: retail picks can look as good as professional ones in some windows (Bradley et al., 2024) without beating the market once risk is considered (Chacon et al., 2023).

Because the data do not include daily prices or a market benchmark, we cannot adjust for risk or for the overall market movement during the window. A single three-month period cannot establish skill, as market direction, sector exposure and event risk can dominate the comparison. These results should therefore be read as descriptive evidence for this window only, and as motivation for the repeated, larger design described in Section 4.4.

7. Limitations

Sample size. Five stocks per group is a demonstration rather than an experiment, and statistical power is very low. **Selection and survivorship.** Rankings are taken at a single date, and forum-mention rankings can be gamed or dominated by a few meme episodes. **Attention is not demand.** Mentions on Reddit indicate interest, not necessarily purchases. **13F**

coverage. 13F filings cover only long positions in US-listed equities above a reporting threshold, exclude short positions, and aggregate across managers with very different strategies. **Single period.** Results depend on one market regime. **Risk.** Raw returns ignore risk. Only cross-sectional dispersion is reported; no benchmark adjustment, daily volatility measure or factor model (for example, Fama–French) is estimated. **Frictions.** Transaction costs, bid–ask spreads and market impact are ignored.

8. Conclusion

This paper sets out a transparent and replicable research design for comparing retail-favoured and institutionally accumulated equities. Its main methodological choices are to form both portfolios on the 13F filing deadline so that information timing is equal, to adjust for the market benchmark, and to use small-sample inference. In the study window, Reddit-favoured stocks earned a higher mean return than institutionally accumulated stocks (13.0% versus 6.8%) but with much greater dispersion, and the difference was not statistically significant. Extending the design across quarters and to larger cross-sections would allow firmer conclusions about whether retail crowds can match professional managers.

References

1. Barber, B. M., & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance*, 55(2), 773–806.
2. Bradley, D., Hanousek, J., Jame, R., & Xiao, Z. (2024). Place your bets? The value of investment research on Reddit’s WallStreetBets. *Review of Financial Studies*, 37(5).
3. Chacon, R. G., Morillon, T. G., & Wang, R. (2023). Will the Reddit rebellion take you to the moon? Evidence from WallStreetBets. *Financial Markets and Portfolio Management*.
4. Eaton, G. W., Green, T. C., Roseman, B., & Wu, Y. (2022). Retail trader sophistication and stock market quality: Evidence from brokerage outages. *Journal of Financial Economics*, 146(2), 502–528.
5. Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2), 383–417.
6. Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *Journal of Finance*, 23(2), 389–416.
7. Li, E. (2022). Alpha in forums: Applying regression analysis to Reddit’s r/WallStreetBets [Senior thesis]. Princeton University.
8. Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, 18(1), 50–60.
9. Pedersen, L. H. (2022). Game on: Social networks and markets. *Journal of Financial Economics*, 146(3), 1097–1119.
10. U.S. Securities and Exchange Commission. (n.d.). Form 13F – Reporting manager holdings. EDGAR. <https://www.sec.gov/edgar>

Appendix A: Data Processing Source Code (C++)

The program below computes the average return and win rate for each group from the prices in Table 2. It is reproduced without modification.

```
#include <iostream>
#include <iomanip>
#include <string>
using namespace std;
int main() {
    string redditNames[5] = {"GME", "AMC", "TSLA", "NVDA", "PLTR"};
    double redditStart[5] = {20, 5, 250, 100, 20};
    double redditEnd[5] = {25, 4.5, 275, 120, 24};

    string institutionalNames[5] = {"AAPL", "MSFT", "GOOGL", "AMZN", "META"};
    double institutionalStart[5] = {180, 400, 150, 180, 500};
    double institutionalEnd[5] = {195, 420, 165, 190, 525};

    double redditTotal = 0;
    double institutionalTotal = 0;
    int redditWinners = 0;
    int institutionalWinners = 0;

    cout << fixed << setprecision(2);

    for (int i = 0; i < 5; i++) {
        double returnPercent = (redditEnd[i] - redditStart[i]) / redditStart[i] * 100;
        redditTotal += returnPercent;
        if (returnPercent > 0) redditWinners++;
    }

    for (int i = 0; i < 5; i++) {
        double returnPercent = (institutionalEnd[i] - institutionalStart[i]) / institutionalStart[i] *
100;
        institutionalTotal += returnPercent;
        if (returnPercent > 0) institutionalWinners++;
    }

    cout << "Reddit average return: " << redditTotal / 5 << "%" << endl;
    cout << "Reddit win rate: " << (double)redditWinners / 5 * 100 << "%" << endl;
    cout << "Institutional average return: " << institutionalTotal / 5 << "%" << endl;
    cout << "Institutional win rate: " << (double)institutionalWinners / 5 * 100 << "%" << endl;
    return 0;
}
```