# Good/Bad Trader Segmentation as a Signal Filter for Crowd-Fade Strategies:
# An Empirical Investigation

*Victory Markets — Proprietary Strategy Research*
*July 2026*



## Abstract

Active Victory Markets strategies generate signal by fading the aggregate net positioning of the retail trading crowd. This research summary examines whether decomposing that crowd into “good” and “bad” traders, defined by realised track record, and trading against the bad cohort specifically, could outperform fading the crowd as a whole. Across four linked tests — a funded-versus-challenge quality comparison, a full-population walk-forward ranking system, a directional follow/fade experiment, and a head-to-head comparison of bad-cohort-only fades against whole-crowd fades — we find no evidence in the tested sample that trader-quality segmentation improves on the existing whole-crowd fade. The historical simulation of the whole-crowd fade produced an excess return of +0.40% per day at a 60% hit rate, while every bad-cohort variant tested (bottom 20%, bottom 30%, bottom 50%, and all net losers) produced only +0.06% to +0.16% per day at hit rates of approximately 52–53%. We argue this result follows from a structural property of the data: bad cohorts represent only 11–20% of total crowd flow, so isolating them discards the aggregate positioning pressure that the fade signal appears to depend on. Within this dataset and methodology, the whole-crowd fade was the best-performing signal among the variants tested; replacing the fade target with a bad-trader-only target would have reduced performance in the tests conducted. We close by outlining three untested ideas that use the good/bad split as a conditional filter rather than a replacement signal, and recommend an agreement-filter test as the next research step.

## 1. Introduction

Victory Markets' current suite of live and funded-account strategies is built around a single core mechanic: fading the net open position (NOP) of the aggregate retail crowd on XAUUSD and related instruments. The premise is that retail order flow, taken in aggregate, tends to be positioned against the prevailing move with sufficient regularity that trading opposite the crowd has produced a historically observed statistical relationship in the tested data. This treats every trader in the flow-generating population identically, regardless of individual track record.

A natural extension is to ask whether not all retail flow is equally informative. If a subset of traders is reliably unprofitable, fading that subset specifically, rather than the crowd as a whole, might sharpen the signal by removing dilution from skilled or neutral participants. This summary tests that hypothesis directly, using broker-level order volume and net-open-position data together with a full realised-P&L trading history for the account population on file.

The question addressed is narrow and practical: does splitting traders into good and bad cohorts by track record, and acting on that split, do better than fading everyone? The remainder of this paper describes the four tests conducted, reports results, explains why the replacement hypothesis fails, and proposes follow-on tests that use the good/bad distinction as a filter rather than a substitute for the existing signal.

This question sits within a broader body of practitioner thinking about retail order flow as a contrarian indicator. The general intuition behind crowd-fade strategies is that retail participants, taken collectively, tend to over-extrapolate recent price action, chase breakouts late, and hold losing positions longer than winning ones — behavioural tendencies that are well documented in the retail-trading literature and that produce systematic mispositioning relative to subsequent price movement. What is less well established, and what motivates the present study, is whether that mispositioning is uniformly distributed across the retail population or concentrated in an identifiable subset of participants. If it is concentrated, isolating that subset should improve the signal; if it is not, isolating any subset should weaken it. The tests below were designed to discriminate between these two possibilities using the account population actually on file, rather than relying on assumptions about where mispositioning originates.

## 2. Methodology

Four linked analyses were conducted in sequence, each building on the last. All tests were run on the full available population of Prop Firm A trading accounts, using incremental FIFO-netted realised P&L and daily XAUUSD price history for forward-return calculation.

The account population spans challenge-phase and funded-phase traders across the observation window, with individual trade records aggregated into daily net positioning figures per trader before ranking or cohort assignment. Realised P&L was computed on a strict first-in-first-out basis so that partial closes and re-entries within a single trader's history are netted consistently, avoiding the distortion that would arise from treating each fill as an independent outcome. Every ranking and every forward-return calculation described below was generated walk-forward, using only information that would have been available at the time, so that none of the reported figures benefit from look-ahead bias.

### 2.1 Funded vs. Challenge Quality Gap

The first step examined whether "trader quality" is measurably differentiated within the account population rather than merely assumed. Funded accounts (traders who have passed evaluation and manage firm capital) were compared against challenge accounts (traders still in evaluation) on win rate and profit factor. Funded traders showed a 55% win rate at PF 1.06; challenge traders showed 36% at PF 0.86 — an observed performance differential within this sample. This comparison is descriptive rather than, on its own, proof of a persistent skill effect. It also establishes a boundary condition: the funded population is a small, self-selected minority, and its observed advantage is not directly harvestable simply via the funded/challenge label, since that label is not available as a live, forward-looking classifier.

### 2.2 Full-Population Walk-Forward Ranking

To classify traders prospectively rather than by the funded/challenge label, a walk-forward ranking tool re-ranks every Prop Firm A trader each day by all-time realised P&L, computed incrementally via FIFO netting with no look-ahead. A recency filter and warm-up period prevented new or dormant accounts from distorting the ranking. Daily XAUUSD price history was used to compute accurate forward returns for all subsequent tests.

### 2.3 Directional Follow/Fade Test

Using the ranking, a directional test followed the top decile's positioning and, separately, faded the bottom decile's. Following the top decile initially returned +33%, but this was concentrated in trending conditions; de-trended, it turned negative at a 45% hit rate. Fading the bottom decile alone produced a 51% hit rate — close to chance in this sample. Neither direction produced a usable standalone edge under the tested methodology.

### 2.4 Fade Bad-Crowd vs. Whole-Crowd (Strategy-Native Test)

The most direct test of the replacement hypothesis applied the existing, unchanged fade logic to two positioning inputs: (a) a bad-cohort-only subset, and (b) the whole crowd, as used in live strategies today. Four bad-cohort definitions were tested — bottom 20%, bottom 30%, bottom 50%, and all net losers — holding the fade mechanism constant so that any performance difference is attributable to the composition of the signal, not to strategy logic.

This design choice — holding the fade mechanism itself fixed and varying only the population being faded — is what makes the test strategy-native rather than purely academic. Earlier tests in this study (Sections 2.1–2.3) established an observable performance difference between trader cohorts and showed that a naive directional signal built on top-decile or bottom-decile positioning does not survive de-trending in the tested sample. Section 2.4 asks a different and more targeted question: given the fade mechanism that is already deployed and was historically positive under the tested methodology, does substituting a narrower, ostensibly higher-quality positioning input into that same mechanism improve on the version already running against the full crowd? Four cutoffs were chosen to span the plausible range of “bad,” from a strict bottom-quintile definition to the broadest possible definition of any account with negative all-time realised P&L, so that the result would not hinge on an arbitrarily chosen threshold.

## 3. Results

The whole-crowd fade outperformed every bad-cohort variant by a wide margin in the historical simulation: +0.40%/day at a 60% hit rate, versus only +0.06% to +0.16%/day at roughly 52–53% hit rates for the bad-cohort variants — materially weaker, and consistent across all four cutoffs tested. Table 1 summarises the comparison.

| Signal tested | Historical excess return | Hit rate | Flow share |
|---|---|---|---|
| Whole-crowd fade — historical simulation using current strategy logic | +0.40%/day | 60% | 100% |
| Bad cohort — bottom 20% fade | +0.16%/day | ~53% | ~11% |

| Signal tested | Historical excess return | Hit rate | Flow share |
|---|---|---|---|
| Bad cohort — bottom 30% fade | +0.12%/day | ~53% | ~14% |
| Bad cohort — bottom 50% fade | +0.08%/day | ~52% | ~18% |
| Bad cohort — all net losers fade | +0.06%/day | ~52% | ~20% |

*Note: flow-share figures are approximate ranges across the tested cutoffs (11–20% of total crowd flow).*

*Performance-basis note: the return and hit-rate figures in this summary are research outputs derived from historical analysis and/or simulation. They should not be interpreted as realised client returns or as net investable returns. This summary does not specify the exact observation dates, number of signal days, or whether spreads, slippage, financing charges, market impact and other transaction costs are included in every reported figure; those items should be documented alongside the underlying research dataset before the results are relied upon for implementation or external comparison.*

## 4. Discussion: Why the Replacement Hypothesis Failed

The bad cohorts, however defined, represent only 11–20% of total crowd flow — the key to why isolating them weakens rather than strengthens the fade signal in this sample. The historical performance of the whole-crowd fade appears to reflect an aggregate-flow effect: it works because it captures the net positioning pressure of the entire retail population, not because any identifiable subgroup is reliably wrong. Stripping the bad-trader subset out and fading it alone therefore removed signal rather than noise in the tests conducted. The excluded 80–89% of flow is not merely dilution sitting on top of a cleaner signal; it appears to be a substantial part of the aggregate pressure the fade responds to — which also explains why performance degrades further as the cutoff narrows from all net losers (~20% of flow) to the bottom 20% (~11% of flow).

This has a useful analogy in market microstructure more generally: a large share of what looks like "noise" in aggregate order flow is not noise from the point of view of price formation — it is the flow. A filter designed to strip out the supposedly uninformed portion of that flow will, almost by construction, also strip out a meaningful share of the pressure that the fade signal is actually responding to, unless the uninformed portion happens to be large relative to the whole. Here it is not: even the broadest bad-cohort definition tested, all net losers, accounts for roughly one-fifth of flow, leaving four-fifths of the signal on the table. The pattern of results across all four cutoffs — monotonically weaker performance as the definition narrows — is consistent with this mechanism and inconsistent with the alternative explanation that a small number of skilled traders are actively pulling the whole-crowd fade off course.

**CONCLUSION**
Within the available dataset and methodology, fading the whole crowd was the best-performing variant tested. Replacing the fade signal with a bad-trader-only signal would have reduced performance in every cutoff examined.

### 4.1 Practical Implications

For live and funded-account deployment, this result argues against any near-term change to the positioning input used by current fade strategies. On the evidence in this sample, engineering effort should not be directed toward building a bad-trader-only feed as a replacement for the existing whole-crowd NOP aggregation, since every cutoff tested produced weaker historical results. The funded-versus-challenge performance gap identified in Section 2.1 remains an interesting property of the account population, but it is a property of trader outcomes rather than a demonstrated standalone trading signal in its current form — the funded label is not observable in advance for a live account and cannot be substituted for the walk-forward ranking used elsewhere in this study.

## 5. Ideas That Could Still Work (Untested)

The tests above evaluated cohort positioning only as a standalone replacement signal. None tested using the good/bad split as a conditional filter on top of the whole-crowd fade that was historically positive in the tested sample — a filter only needs to carry conditional information, not reconstruct the flow-share dominance the whole-crowd signal already provides. This is an important distinction: a replacement signal must be self-sufficient, generating its own edge from its own positioning input, whereas a filter is only ever asked to answer a narrower conditional question — is the existing signal more or less reliable under a particular condition — and can therefore add value even when built on a comparatively weak underlying cohort. Three candidate approaches follow, in increasing order of implementation complexity.

- Agreement filter. Fade the whole crowd as at present, but check the better-performing trader cohort: if positioned toward the fade, take full or elevated size; if aligned with the crowd instead, skip the trade or reduce size. This uses the better-performing cohort as a veto or confidence layer rather than as a primary signal, and is the simplest of the three to implement since it requires only a same-day comparison of two already-computed positioning aggregates.
- Divergence signal. Days where good and bad traders are positioned in opposite directions — the bad crowd leaning one way, good traders leaning the other — may represent the highest-conviction fade opportunities, since they identify moments where informed and uninformed flow are actively disagreeing rather than moving together. This condition has not yet been measured in isolation and would require classifying every trading day by the sign of both cohorts' aggregate positioning before any performance comparison can be run.
- Conviction weighting. Size the fade continuously by how lopsided the lower-performing cohort's positioning is relative to the better-performing cohort's, as a proxy for the degree of behavioural imbalance at a given moment. This is the most complex of the three, since it moves from a binary filter to a continuous sizing function and would need its own calibration, robustness testing and backtest before any live use.

## 6. Limitations and Caveats

**CAVEAT**

The “good” traders here are only weakly good (PF ~1.06, small share of flow), so a filter built on them may add little. This is a hypothesis to test, not a conclusion.

All forward-return figures derive from XAUUSD price history over the sample period and the specific Prop Firm A account population on file; results should be re-validated before generalising to other symbols, brokers, or periods. The walk-forward ranking and cohort assignments used throughout are also sensitive to the recency filter and warm-up parameters chosen for this study; a materially different warm-up window could shift accounts between cohorts near the boundary and should be checked as a robustness step before any of the untested ideas in Section 5 are taken further. Finally, the profit-factor differential underlying the good/bad split (1.06 versus 0.86) is modest in absolute terms, and the population used is drawn from a single prop-firm partnership; broader validation across additional funded-account partnerships would strengthen confidence that the pattern observed here is general rather than specific to this dataset.

The reported performance statistics should therefore be read as comparative research metrics within the tested framework, not as a forecast, promise or representation of future returns. Before any strategy change is implemented, the results should be reproduced with a fully specified observation window, signal count, execution assumptions, transaction-cost treatment and appropriate robustness checks.

## 7. Recommended Next Step

Test the agreement filter first: split existing fade days into “good traders agree with the fade” versus “good traders oppose it,” and compare forward returns. A clear advantage when good traders are aligned marks a usable filter worth developing toward live deployment; identical buckets mean the good/bad data adds nothing conditionally, and the line can be closed.

This next step is deliberately the cheapest of the three untested ideas to run: it requires no new cohort definitions beyond those already built for this study, no continuous sizing calibration, and no new data collection — only a same-day comparison of two already-available positioning aggregates against forward returns already computed for the whole-crowd fade. A clear result either way would resolve the open question left by this research thread with a minimal additional testing burden.

## 8. Conclusion

This study set out to test a specific, practical hypothesis: that fading a bad-trader subset of the retail crowd would outperform fading the crowd as a whole. Across four linked tests — a funded-versus-challenge quality comparison, a full-population walk-forward ranking, a directional follow/fade experiment, and a strategy-native comparison of bad-cohort and whole-crowd fades — the hypothesis is not supported in the tested sample. The historical simulation of the current whole-crowd fade logic materially outperformed every tested bad-cohort variant, and the pattern appears structural rather than incidental: the bad cohorts, however defined, never account for more than roughly a fifth of total crowd flow, so isolating them discards most of the aggregate positioning pressure that the fade signal appears to depend on. The observed performance gap between funded and challenge accounts remains noteworthy, but it is not, in its current form, a substitute for the existing signal. The more promising direction for further research is not replacement but conditioning: using the good/bad split as a filter or confidence layer on top of the whole-crowd fade, beginning with the agreement-filter test recommended above.

## IMPORTANT INFORMATION, REGULATORY DISCLOSURE AND RISK WARNING

**Regulatory status.** Victory Technologies Ltd, trading as Victory Markets (“Victory Markets”), is licensed by the Financial Services Commission of Mauritius as an Investment Dealer (Full Service Dealer, excluding Underwriting), Licence No. GB25204886. Victory Markets’ regulatory status should not be understood as approval, endorsement or verification by the Financial Services Commission of Mauritius of this document, the research methodology described herein, any trading strategy, or any performance information or conclusion contained herein.

**Nature and purpose of this research.** This document constitutes proprietary strategy research prepared for general informational and research-discussion purposes only. It is not independent investment research, investment advice, a personal recommendation, an offer or solicitation to buy or sell any financial instrument, or an invitation or recommendation to participate in any particular investment or trading strategy. Nothing in this document has been prepared having regard to the investment objectives, financial circumstances, knowledge, experience or particular needs of any specific recipient. Recipients should make their own independent assessment of the information and, where appropriate, obtain independent financial, legal, tax, accounting and regulatory advice before making any investment or trading decision.

**Historical, simulated and backtested results.** Unless expressly identified otherwise, performance figures, hit rates, forward returns, profit factors and other statistics contained in this document are derived from historical analysis and/or simulated or backtested strategy results using the dataset, methodology and assumptions described in the research. Such results do not represent a promise, forecast or guarantee of future performance and should not be interpreted as actual returns earned by any particular investor or client. Hypothetical and backtested results have inherent limitations and may benefit from assumptions, modelling choices or market conditions that will not recur in live trading. Actual trading results may differ materially.

**Performance and costs.** Past performance, historical relationships and simulated results are not reliable indicators of future performance. Unless expressly stated, reported research results should not be assumed to include all transaction costs, spreads, slippage, financing charges, liquidity effects, market impact, taxes or other costs that may arise in actual trading. These factors can materially reduce or eliminate returns indicated by historical analysis or simulation.

**Research methodology and limitations.** The analysis is based upon proprietary, aggregated and/or anonymised trading data together with market data and assumptions described in the document. Reasonable care has been taken in preparing the research and in seeking to ensure that information is accurate and current as at the stated research date. Quantitative results are nevertheless inherently dependent upon the quality and completeness of the underlying data, the observation period, sample population, classification methodology, execution assumptions, parameter choices and prevailing market regime. Results obtained from another period, instrument, population or methodology may differ materially. Conclusions should therefore be understood as applying to the specific tests and data described and should not be generalised beyond them without further validation.

**Conflicts of interest and proprietary trading.** Victory Markets conducts proprietary and institutional trading and execution activities and may hold, establish, modify or close positions in financial instruments discussed in its research, including instruments to which the strategies described herein relate. Victory Markets may itself use research, models, signals or strategies described in this document before, at the time of, or after publication. Victory Markets also has or may have commercial relationships with proprietary trading firms, brokers, liquidity providers, data providers, capital allocators and other market participants whose activities or data may be relevant to the research. Accordingly, this document should not be regarded as research produced independently of Victory Markets’ proprietary or commercial interests.

**Data and confidentiality.** Research may incorporate aggregated or anonymised data supplied by commercial counterparties. The use of such data is intended to analyse population-level behaviour rather than identify or assess any individual trader. No disclosure of confidential counterparty or individual-account information is intended, and the publication or distribution of research remains subject to Victory Markets’ contractual data-use and confidentiality obligations.

**Trading risk.** Trading financial markets, particularly leveraged instruments and contracts for difference (“CFDs”), involves substantial risk and can result in the loss of some or all capital committed. Historical statistical relationships may cease to exist without notice, and strategies that have performed positively in one market regime may experience significant losses or drawdowns in another. No statement in this document should be interpreted as an assurance that losses will be limited or that any strategy will achieve a particular return.

**Distribution, changes in views and responsibility.** This document may only be distributed or used where permitted by applicable law, regulation and contractual obligations and is not directed at any person or jurisdiction where such distribution, use or related activity would be unlawful. The opinions, interpretations and conclusions expressed are those of Victory Markets as at the date of publication and may be revised as additional information, market conditions or research