2026

Algorithmic Trading in Idiosyncratic-Payoff Markets: A Multi-Agent System for On-Chain Prediction Contracts

Saif Aldeen A.K. Agha

Claremont McKenna College

Algorithmic Trading in Idiosyncratic-Payoff Markets:

A Multi-Agent System for On-Chain Prediction Contracts

Submitted to

Professor Fan Yu

by

Saif Aldeen A.K. Agha

for

Senior Thesis

Spring 2026

April 27, 2026

Acknowledgments

I am grateful to Professor Fan Yu for trusting me with this project. I had him as my professor for derivatives during my sophomore year, and his teaching in that class was the reason I knew, well before this thesis began, that I wanted him as my reader. I thank Professor Laura Grant, who led the thesis seminar that kept this work on target through every stage. I thank my academic advisor, Professor Ricardo Fernholz, for the guidance and breadth of perspective he has brought across four years.

I thank the faculty of the Robert Day School of Economics and Finance, who trained me to think the way this thesis demanded. Professor Yaron Raviv made me an academic again. His game-theory course was one of the most important classes I took at CMC; it reinvigorated my love for the work this thesis is built on, and his way of teaching strategic thinking will travel with me long after graduation. Professor Peter Kelly's conversations about poker and niche investment opportunities are among my favorites at the Robert Day School; I am grateful for his friendship. Professor Angela Vossmeyer's econometrics course is the methodological foundation of the empirical strategy in this thesis; I am grateful for her care and clarity. Professor Marc Massoud, who called every student his kid and whose last course at CMC I was privileged to take, felt like family to me.

I thank Professor Richard Burdekin, whose way with words and gift for elucidating complicated ideas remain for me a model of what good communication looks like; Professor Eric Hughson, who helped me understand idiosyncratic risk and the microstructure foundations on which this thesis depends; Professor Nishant Dass, for his friendship; Professor Manfred Keil, whose classes were hard, in the way that matters; Professor Murat Binay, for his corporate-finance instruction; Professor Daniel Firoozi, for the microeconomic foundation he laid; and Professor Laura Dannhäuser, for the kindness she has shown me through her leadership courses.

I thank my partner, Chung ("Kieran") Ching Chi, who joined this project after reading an early draft of this thesis. The work documented here is the foundation for the research program we are building together, and I am grateful to have him on that journey.

I thank Avi Thaker, who took me on as an intern at Tauroi Technologies and taught me a great deal during my time there. I appreciate his continued support and guidance, and I am grateful for everything I learned from him. I am also grateful to Dr. Yazan Al Hasan, who generously contributed compute resources that supported parts of this work.

I thank Nicholas Bagatelos, CMC class of 1986 and my partner on Net Zero Envelope, for showing me what an entrepreneurial life can look like. I thank Jerry Bien-Willner, Claremont McKenna class of 1997, who is a great role model and an example of good character, and Jacob Bain, Claremont McKenna class of 2020, who talked me through my application to the school and through every stage of my career since.

I thank Jerry Fielding, who introduced me to Bitcoin and the broader crypto landscape in 2020. The conversation that began that year is the substrate of nearly everything in this thesis and much of what I have pursued since. I thank Jordan Rose, who knew I was interested in crypto and gave me real work to do that let me explore the space and learn from the inside.

I thank all of CT, especially ITT and Remilia.

I thank Tariq Amin, Claremont McKenna class of 2024.

To my father, Ayad, who left Iraq alone at fourteen years old in 1981, fleeing war, and built a life in this country that culminated in becoming a physician who saves lives every week: he is my model of conviction and of what work can do against impossible odds. When I was entrenched in on-chain work and unable to see past it, he was the one who told me to take artificial intelligence seriously and to think about where my effort would matter most. The thesis you are reading sits at the intersection

of those two suggestions and would not exist without his redirection. To my mother, Jeni, who grew up in Newark, Ohio without the means for college herself, who put my father through medical school working as a corporate recruiter, who runs marathons because that is the kind of person she is: she is my rock, and the spirit she gave me to go forth is the engine of every project I have ever undertaken. I love them.

To my brother, Reyam Jadd Agha, who is taller than me, smarter than me, funnier, better-looking, the best athlete in the family, and whose future I cannot wait to see unfold: I love you. To my sister, Iya Agha, my older sister and second mother, now a dermatologist whose example I have spent my life trying to live up to: I love you, and I am proud of how far you have come.

To my cousins Leith Ghani, who introduced me to public equity markets when I was a freshman in high school, and Zade Ghani, whose advice has been a steady ground throughout: they are the kind of family one is unspeakably lucky to have.

To my grandparents, my Nana Karan and Papa Thurman on my mother's side, my Bebe Souad and Jidu Kamal on my father's side: the lives you built are the ground ours stand on. Thank you.

And to myself: only we really know the work we put in.

Any errors that remain are my own.

Questions about this thesis can be directed to `sagha26@cmc.edu` or `ebk@ebk.tech`.

Abstract

This thesis documents the design, deployment, and forward-test evaluation of an evolutionary multi-agent algorithmic trading system on Polymarket, the largest on-chain prediction market. The system pairs a locally-hosted, refusal-modified 72-billion-parameter language model with a 41-feature gradient-boosted entry filter, an adaptive exit engine, and a continuous-truncation evolutionary selection mechanism that maintains a population of approximately 500 autonomous agents. The central empirical exercise estimates a two-way fixed-effects regression of trade-level profit on the absolute divergence between the agent's calibrated language-model probability estimate and the prevailing market price. On 113,346 resolved trades placed by 3,157 unique agents over a 34-day window, the divergence coefficient is \$245 per trade ($t = 24.5$); a 1,000-iteration permutation test in which divergence is randomly reassigned within each agent fails to recover the observed coefficient in any iteration, placing it $z = 22.3$ standard deviations outside the placebo distribution. Performance persistence is strong: the rank correlation of agent mean profit-and-loss across the first and second halves of each agent's trading history is 0.66, and 81 percent of top-decile agents in the first half remain in the top decile in the second. Cumulative simulated profits over the window total \$24.0 million on an initial simulated capital base of \$1 million; this number is reported as suggestive context rather than as the principal empirical claim, since 83 percent of the cumulative is generated by strategies that were subsequently banned from live deployment for catastrophic underperformance. The placebo-identified divergence coefficient is the result that survives every restriction.

A parallel live deployment on real USDC across 65 wallets (1,791 trades, 19 days) realizes a -7.2% cumulative *realized* return on initial deposits (mark-to-market net asset value approximately offsetting at the experimental cutoff; see Section 9.1), dominated by an operational-risk event in the BUY_YES signal-construction path that placed Kelly-sized positions when the language model had produced no directional estimate. Following the documented engineering fix on April 20, 2026, the post-fix sample (784 trades, 8 days) yields an 82.5 percent win rate and positive realized profit; this post-fix sample is reported as a small-n descriptive directional confirmation rather than as a free-standing performance estimate. The asset class itself avoids the systematic factor exposure central to the Duarte, Longstaff and Yu (2007) characterization of fixed-income arbitrage. Prediction-market contracts resolve to idiosyncratic events; residual systematic risk is concentrated in the platform and execution layers, and that relocation of risk is itself a substantive feature of the new asset class.

Keywords. prediction markets; algorithmic trading; large language models; evolutionary algorithms; operational risk; decentralized finance; Polymarket.

JEL classification. G12, G14, G17, C45, C53.

Table of Contents

Acknowledgments ii

Abstract iv

1 Introduction 1

2 Related Literature 2
- 2.1 Prediction-market efficiency . . . 2
- 2.2 Machine learning and language models in trading . . . 3
- 2.3 Defaultable claims and operational risk . . . 4
- 2.4 Position of the present paper . . . 5

3 Institutional Setting 5
- 3.1 Polymarket . . . 5
- 3.2 The credit-instrument analogy . . . 5
- 3.3 Order flow and transaction costs . . . 6

4 System Architecture 6
- 4.1 Compute and inference stack . . . 6
- 4.2 Base model and modification . . . 7
- 4.3 Inference pipeline . . . 7
- 4.4 Quant pipeline and position sizing . . . 8
- 4.5 Entry and exit gates . . . 8
- 4.6 Evolutionary selection . . . 9
- 4.7 Strategy primitives . . . 10

5 A Simple Theoretical Framework 11
- 5.1 Setup . . . 11
- 5.2 Market clearing . . . 11
- 5.3 Algorithmic-trader profitability . . . 11
- 5.4 Implications for the empirical specification . . . 12

6 Data 12
- 6.1 Sample construction . . . 12
- 6.2 Variables . . . 13
- 6.3 Summary statistics . . . 14
- 6.4 Sample composition . . . 14

7 Empirical Strategy 15
- 7.1 Main specification . . . 15
- 7.2 Identification . . . 16
- 7.3 Placebo test . . . 16

8 Paper-Sample Results 17
- 8.1 Main regression . . . 17

8.2 Permutation placebo . . . 18
8.3 Heterogeneity by market category . . . 18
8.4 Divergence buckets . . . 19
8.5 Daily P&L . . . 20
8.6 Risk-adjusted return . . . 21
8.7 Robustness . . . 22
8.7.1 Restricted to live-surviving strategies . . . 22
8.7.2 Stability across the experimental window . . . 23
8.7.3 Top-trade concentration . . . 23
8.7.4 Price-extremity control . . . 24
8.8 Performance persistence . . . 25
8.9 Caveats on the headline numbers . . . 25
9 Live Deployment 27
9.1 Live sample . . . 27
9.2 Direction asymmetry and the operational-risk event . . . 27
9.3 Diagnosing the BUY_YES failure . . . 28
9.4 Operational-risk interpretation . . . 30
9.5 Other live observations . . . 32
9.6 Live risk-adjusted return . . . 33
10 Discussion 33
10.1 Interpretation . . . 33
10.2 Limitations . . . 34
10.3 Implications and extensions . . . 35
11 Conclusion 36

1 Introduction

A binary prediction-market contract is a defaultable claim with zero recovery: it pays one dollar if a specified event occurs and zero otherwise, with a fixed resolution date and a venue-specific dispute-resolution mechanism. The mathematical structure of these instruments is identical to the binary recovery component of a defaultable corporate claim, and their pricing is subject to the same informational and behavioral frictions that the credit-derivatives literature has documented over the past two decades (Jarrow and Yu, 2001). What is new is the venue, the volume, and the participant mix. As of early 2026, the largest decentralized prediction-market protocol intermediates volume in the billions of dollars across thousands of contracts spanning U.S. and international elections, sports outcomes, cryptocurrency price thresholds, and geopolitical events; major news organizations now cite contract prices alongside conventional polling.

This thesis asks whether an algorithmic trading system, built around a frontier language model and an evolutionary multi-agent architecture, can systematically identify and exploit mispricing in such contracts. The question has both an empirical and a methodological component. Empirically, the question is whether a well-designed automated counterparty earns positive risk-adjusted returns against a participant pool dominated by retail traders subject to the probability-weighting biases formalized by Kahneman and Tversky (1979) and Prelec (1998). Methodologically, the question is what the relevant unit of analysis is when alpha is generated by a population of agents whose composition itself is shaped by an evolutionary selection mechanism. Conventional algorithmic-trading research analyzes a single strategy with fixed parameters; the system studied here is closer in spirit to an adaptive market-making operation than to a one-shot arbitrage.

The paper makes four contributions.[1] First, it documents the design and live deployment of a system that pairs a refusal-modified open-weights 72B language model, locally hosted on a single GPU server, with a 41-feature gradient-boosted entry filter and a continuous-truncation evolutionary selection mechanism over a population of approximately 500 autonomous agents. The architecture is described in sufficient detail to make the experimental procedure reproducible in spirit, while specific parameter weights and proprietary signal mixes are omitted to preserve commercial sensitivity. Second, the paper assembles what is, to the best of my knowledge, the first trade-level panel of language-model-driven prediction-market trades with matched signal provenance, agent identifiers, divergence measurements, and settlement outcomes: 113,346 resolved paper trades over 34 days plus 1,791 live trades with real on-chain capital. Third, the paper estimates the within-agent relationship between model-market divergence and trade-level profit and shows that a one-unit increase in absolute divergence is associated with approximately $245 of additional per-trade profit, with the result robust to a 1,000-iteration permutation test that randomly reassigns divergence values within each agent. Fourth, the paper documents a clean experimental decomposition of paper-versus-live per-

[1] The system described in this paper is the proprietary work of EBK LLC; the design, implementation, empirical analysis, and writing in this thesis are the author's own.

formance divergence: the live deployment's negative realized P&L is attributable entirely to a documented operational-risk event in the execution layer (a Kelly-direction inversion affecting the BUY_YES side), with the post-fix sample showing positive returns at an 82.5 percent win rate over a small sample window.

A central interpretive claim runs through the empirical analysis. The Duarte et al. (2007) characterization of fixed-income arbitrage strategies as *nickels in front of a steamroller*, strategies that earn modest positive carry under normal conditions but suffer catastrophic losses under correlated stress, depends on the existence of common-factor exposure across positions in the strategy book. Prediction-market contracts resolve to idiosyncratic events: the outcome of a specific basketball game, the closing price of a specific cryptocurrency on a specific date, the result of a specific election. A portfolio diversified across categories has no analogue to the duration, credit-spread, or liquidity factors that drive correlated arbitrage losses in conventional fixed-income markets. The conventional steamroller is absent.

This does not mean tail risk disappears. The platform itself carries risk (oracle disputes, smart-contract failure, regional regulatory action), and the execution layer carries operational risk of the kind documented by Chernobai, Jorion and Yu (2011) for U.S. financial institutions. The live drawdown analyzed in Section 9 is precisely an operational-risk realization in this sense: a documented engineering bug in the order-routing layer caused a structural loss on a specific trade direction, was identified and fixed within the experimental window, and post-fix performance reverted to the pattern predicted by the paper-sample analysis. The asset class moves systematic risk *out* of the underlying contracts and *into* the infrastructure that intermediates them. Whether this is a net improvement depends on whether one trusts the infrastructure more than one trusts correlated credit markets, an empirical question outside the scope of this thesis but not outside its policy relevance.

The remainder of the paper proceeds as follows. Section 2 positions the paper in three relevant literatures: prediction-market efficiency, machine learning in algorithmic trading, and the credit-derivatives literature on defaultable claims and operational risk. Section 3 describes Polymarket as a venue and draws the explicit credit-instrument analogy. Section 4 documents the system architecture: the local inference stack, the agent population, the evolutionary selection mechanism, and the gradient-boosted entry and exit filters. Section 5 presents a simple theoretical framework with behaviorally biased traders and calibrated algorithmic traders. Section 6 describes the paper and live samples. Section 7 specifies the empirical strategy. Section 8 reports paper-sample results. Section 9 reports live-sample results and decomposes the operational-risk drawdown. Section 10 discusses limitations and extensions.

2 Related Literature

2.1 Prediction-market efficiency

Wolfers and Zitzewitz (2004) establish that prediction-market prices track realized event frequencies closely and frequently outperform structured polls and expert forecasts; they treat prediction markets as an information-aggregation technology and evaluate them, at least ap-

proximately, against a semi-strong-form efficiency benchmark. The methodological foundations are due to Hanson (1995), who proposed logarithmic-scoring market mechanisms as a way to incentivize truthful probability reporting, and to the long-running Iowa Electronic Markets program documented across two decades of election-futures and other research (Berg, Forsythe, Nelson and Rietz, 2008b; Berg, Nelson and Rietz, 2008a), which established that prediction markets can outperform polls in real-money settings. Manski (2006) shows that under heterogeneous beliefs and budget constraints, market-clearing prices can deviate systematically from the mean subjective probability of the trading population, with the magnitude of the deviation an increasing function of belief dispersion. Snowberg and Wolfers (2010) document and decompose the favorite-longshot bias, attributing it to a mixture of risk preferences and probability misperception. Hahn and Tetlock (2006) review the policy-relevance literature on information markets. Subsequent empirical work documents calibration failures at the tails of the probability distribution: long shots are systematically overpriced and heavy favorites systematically underpriced relative to realized frequencies. The existence of these systematic miscalibrations creates potential profits for an informed counterparty. This thesis tests, directly, whether such an informed counterparty exists in the form of a language-model-driven agent system.

2.2 Machine learning and language models in trading

The earlier wave of machine-learning-in-trading research focuses on supervised and reinforcement-learning architectures. Deng, Bao, Kong, Ren and Dai (2017) apply deep reinforcement learning to high-frequency stock and futures trading and demonstrate that policy-gradient methods can learn profitable strategies. Fischer and Krauss (2018) show that long short-term memory networks generate economically significant returns in a statistical-arbitrage context on S&P 500 constituents. Sutton and Barto (2018) provide the canonical theoretical treatment of reinforcement learning.

A more recent literature substitutes the language model for these bespoke architectures. Lopez-Lira and Tang (2023) use a frontier language model to generate sentiment scores on financial news and find that a long-short portfolio constructed from those scores earns positive risk-adjusted returns; the result motivated a wave of LLM-based prediction-and-trading work synthesized in two recent surveys. Yan, Aste and Zhao (2025) trace the trajectory from deep-learning alpha pipelines to language-model-based pipelines and identify the central methodological shift as a move from end-to-end learned representations to heterogeneous, text-grounded signal extraction. Li, Hou, Wang, Yang and Yang (2025) provide the most comprehensive contemporary taxonomy of language-model methods in financial prediction, organizing the field into sentiment extraction, point-in-time information extraction, numerical reasoning, multimodal signal generation, agentic execution, and governance functions. Both surveys explicitly identify multi-agent LLM trading systems as an emerging frontier.

Within that frontier, two closely-related works are immediately relevant. Xiong, Zhang, Feng, Sun and You (2025) introduce *QuantAgent*, a price-driven multi-agent LLM framework for high-frequency trading on equities and cryptocurrency futures, with specialized indicator,

pattern, trend, and risk agents that exchange structured reasoning. The framework demonstrates that decomposing a trading problem across coordinated LLM agents can outperform monolithic single-agent baselines on conventional return-prediction benchmarks. Darmanin (2025) pairs a language model with a reinforcement-learning agent in a contemporaneous M.Sc. thesis and shows that LLM-generated strategy guidance can shape the risk profile of a single underlying RL policy across distinct investor mandates without retraining the policy itself.

The current paper differs from this body of work in three respects. First, the asset class is binary prediction-market contracts rather than equities, futures, or cryptocurrency, with a payoff structure that maps cleanly onto the defaultable-claim framework of the credit-derivatives literature (Section 3). Second, the multi-agent architecture in this paper is *evolutionary* rather than functionally specialized: agents share a common pipeline and are differentiated by their genome (a vector of strategy parameters), with selection pressure operating over a continuously-replenished population. The functional decomposition of Xiong et al. (2025) is orthogonal to, and could in principle compose with, the evolutionary mechanism studied here. Third, the paper presents a matched-pair paper-versus-live evaluation in which the live deployment operates with real on-chain capital, a forward test of the kind that Li et al. (2025) flag as conspicuously underrepresented in the existing FinLLM literature, where simulation studies dominate.

2.3 Defaultable claims and operational risk

The credit-derivatives literature provides the closest structural analogue to prediction-market contracts. Jarrow and Yu (2001) develop the foundational framework for pricing defaultable securities under counterparty risk. Yu (2005) documents the role of accounting transparency in the term structure of credit spreads. Yu (2006) empirically evaluates a sophisticated arbitrage strategy (capital structure arbitrage) and shows that simulated returns substantially overstate realizable returns once execution and liquidity costs are imposed; a pattern that this paper finds in a new venue. Duarte et al. (2007) characterize fixed-income arbitrage strategies as "nickels in front of a steamroller": positive carry under normal conditions, catastrophic losses under correlated stress. Qiu and Yu (2012) document how liquidity in credit derivatives is endogenously thinnest precisely when traders need it most, a pattern with clear analogue in the on-chain order book during contract resolution.

A separate strand within this literature studies *operational risk*: losses arising not from market movements but from technology failures, process errors, and human mistakes. Chernobai et al. (2011) document the empirical determinants of operational risk in U.S. financial institutions and show that operational losses are a first-order component of realized returns in algorithmic strategies. The high-frequency microstructure literature provides the natural complement: Easley, López de Prado and O'Hara (2012) formalize order-flow toxicity in continuous-auction markets and show that adverse-selection costs are a major driver of realized strategy underperformance, and López de Prado (2018) synthesizes the broader integration of microstructure and machine-learning methods into the algorithmic-trading literature. Section 9

of this paper documents an operational-risk event in precisely the Chernobai et al. (2011) sense: a documented engineering bug in the order-routing layer that caused systematic losses on a specific trade direction, was identified and fixed within the experimental window, and is clearly attributable to implementation rather than to the underlying signal.

2.4 Position of the present paper

Against this background, the contribution can be stated concisely. Prediction markets are broadly but imperfectly efficient, with miscalibration concentrated in the tails of the probability distribution. The contemporary FinLLM literature has begun to construct multi-agent language-model trading systems but has, to my knowledge, neither tested this architecture in prediction markets nor evaluated it in a live deployment with on-chain capital. The trade-level panel assembled here permits the required test; the agent fixed-effects specification cleanly separates within-agent signal quality from between-agent selection effects; the live deployment provides the forward test that simulation-only studies cannot; and the operational-risk decomposition of the live drawdown isolates signal generation from execution in a way that algorithmic-trading research more typically elides.

3 Institutional Setting

3.1 Polymarket

Polymarket is a decentralized prediction-market protocol deployed on the Polygon proof-of-stake blockchain. Users deposit USDC, a dollar-pegged stablecoin, into non-custodial wallets and place orders through a central limit order book operated by the protocol. Each market consists of a pair of binary contracts, YES and NO, whose prices are bounded between zero and one dollar and which sum to one dollar under no-arbitrage. When the underlying event resolves, the contract corresponding to the realized outcome pays one dollar and the other pays zero. Resolution is handled by UMA's optimistic oracle, which permits a fixed challenge window during which any participant may dispute the proposed outcome by posting a bond.

At the time of writing, Polymarket intermediates volume in the low billions of dollars per month across thousands of active contracts. Market categories include U.S. and international elections, cryptocurrency price milestones, sports outcomes (NFL, NBA, MLB, NCAA, Premier League, esports), weekly and monthly macroeconomic releases, and a residual category covering entertainment, science, and miscellaneous events. Market lifespan ranges from minutes (for intra-game sports contracts) to a year or more (for long-horizon political and macroeconomic contracts).

3.2 The credit-instrument analogy

A binary prediction-market contract has the same payoff structure as a zero-recovery defaultable claim. Let $\tau \in [0, T]$ denote the random resolution time of a market with terminal date T, and let $X \in \{0, 1\}$ denote the binary realization at τ. The YES contract pays $\mathbf{1}\{X = 1\}$ and the NO contract pays $\mathbf{1}\{X = 0\}$ at τ. This structure is mathematically identical to the binary recovery component of a defaultable corporate claim with zero recovery on default, where

"default" is replaced by "the event-of-interest realizes." The pricing problem is correspondingly the same: identify the risk-neutral probability of the binary event, conditional on the available information, and compare to the market price.

The analogy is more than cosmetic. The literature on defaultable claims (Jarrow and Yu, 2001, Yu, 2005, Yu, 2007) provides the relevant pricing framework, including the treatment of counterparty risk. In the prediction-market context, counterparty risk takes a specific form: a successful UMA oracle dispute that overturns a proposed resolution outcome, or a smart-contract failure that prevents settlement. These risks are platform-level rather than counterparty-level in the conventional credit sense, but their effect on the pricing of the binary claim is analogous. This thesis treats Polymarket contracts as a new asset class within the broader category of defaultable binary instruments, and inherits the analytical machinery developed for that broader class.

3.3 Order flow and transaction costs

Trades execute through a central limit order book that matches incoming limit and market orders against a resting order queue. Bid-ask spread is the primary explicit transaction cost for liquidity takers and varies systematically with contract price, market category, and time to resolution. Contracts priced near the extremes of the unit interval (below 0.10 or above 0.90) carry relatively wider spreads as a percentage of token price; this matters for the execution-cost analysis in Section 9. The protocol charges no explicit trading fee, but on-chain gas costs for order submission, cancellation, and redemption are borne by the trader.

The system described in this paper interacts with Polymarket through the public REST and websocket APIs. Orders are placed as EIP-712 signed messages generated by each wallet's private key; the system controls these keys directly through a custody layer that is segregated from the personal funds of the operator. All order, fill, and settlement data are recorded in a private relational database with contract identifiers that permit later merging against the public on-chain settlement record.

4 System Architecture

This section describes the system as deployed during the experimental window. The architecture is presented at a level of detail sufficient to make the experimental procedure reproducible in spirit, while specific signal weights, mutation rates, and proprietary parameter values are omitted to preserve commercial sensitivity. Where parameters affect the interpretation of the empirical results, they are reported either in this section or in Section 7.

4.1 Compute and inference stack

The system runs across three classes of compute. A controller node, a workstation-class machine running a Unix-like operating system with 64 GB of memory, hosts the orchestration layer: the paper-mode trading engine, the live-mode trading engine, the auto-maintenance daemon, the redemption monitor, the dashboard generators, and the agent-population manager. Approximately thirty long-lived background processes run continuously under

operating-system service supervision.

The primary inference workload runs on a dedicated GPU server with 128 GB of unified memory, accessed over an encrypted private-network tunnel. Inference is served through vLLM, an open-source high-throughput inference framework that exposes an OpenAI-compatible HTTP API. A health probe paired with a synthetic end-to-end completion check guards against the silent-failure mode in which a naïve liveness check returns success on a tiny payload while real inference hangs.

A tiered failover hierarchy ensures availability: when the primary inference server is unreachable, the system falls back through a hierarchy of secondary inference paths, terminating in a machine-learning-only mode that rejects all entries when no language-model inference is available.

4.2 Base model and modification

The primary inference model is an open-weights 72-billion-parameter transformer in the Qwen-2.5 family, served in 5-bit quantized format (GGUF Q5_K_M) for memory efficiency. The choice of an open-weights model is substantive rather than incidental: closed-API inference at the volume the system requires would impose prohibitive marginal costs and would create a dependency on a third party whose pricing, latency, and refusal policies are outside the operator's control.

The model has been modified from its base-instruct release using the *abliteration* technique (Labonne, 2024), which suppresses the residual-stream refusal direction in selected attention and MLP layers without retraining the model. The motivation is domain-specific: many prediction-market contracts concern human-life or geopolitical outcomes (probability of military action, probability of sovereign default, mortality contracts) on which the base model's alignment training causes systematic refusal or hedged probabilistic output. The abliteration removes this refusal behavior while preserving the model's broader capability profile. The modification is documented in the public open-weights community and is not a bypass of safety properties unrelated to the trading domain.

4.3 Inference pipeline

For a given market, the agent's pipeline executes as follows. The agent receives a structured prompt containing the contract question, the current market price, the order-book depth at top of book, the contract's resolution horizon, the agent's own configured strategy parameters, and a recent slice of relevant news from the system's news ingestion pipeline. The model returns a strict JSON object with four fields: a probability estimate $p_{\text{model}} \in [0, 1]$, a confidence value, a directional recommendation (BUY_YES, BUY_NO, or ABSTAIN), and a free-text reasoning trace. Each agent makes a one-shot call; consensus across agents on the same market is computed downstream by a separate aggregation module that takes a trimmed mean in logit space with a disagreement-based weighting penalty.

Raw probability estimates are passed through a Platt-scaling calibration (Platt, 1999) fitted to a held-out historical sample of agent-emitted probabilities and realized outcomes. The

training-test split is constructed in time order, with a cutoff of April 1, 2026 chosen to leave the test window fully out-of-sample relative to the trades on which the calibrator is fit; this respects the temporal structure of the data and avoids the implicit look-ahead bias of a random split. The fit requires a minimum of 50 training pairs and 20 test pairs to deploy a calibrator. Three calibration families are fit on the training window, namely Platt scaling, beta calibration (Kull, Silva Filho and Flach, 2017), and temperature scaling, and the family that minimizes expected calibration error on the held-out test set is selected for production use. For the primary 72-billion-parameter inference model, the deployed Platt calibrator takes the form $\sigma(A \cdot \text{logit}(p) + B)$ with $A = 0.877$ and $B = 1.621$, reducing test-window expected calibration error from 0.252 (raw model output) to 0.145 (calibrated output) and the Brier score from 0.259 to 0.211. The improvements meaningfully affect downstream Kelly sizing because Kelly position size is highly sensitive to estimation error in $\hat{p}$ at extreme market prices: at $p_{\text{market}} = 0.90$, a five-percentage-point miscalibration would produce a fifty-percent overbet under the unshrunk full-Kelly rule. Calibrated probabilities are clipped to $[0.001, 0.999]$ before entering the position-sizing module described in the next subsection.

4.4 Quant pipeline and position sizing

Given a calibrated probability estimate and a market price p_{market}, the system computes an edge measure $e = \tilde{p}_{\text{model}} - p_{\text{market}}$. Position sizing follows a fractional-Kelly rule (Kelly, Jr., 1956): the recommended bet size is a strategy-specific fraction of the full-Kelly fraction, with the fraction shrunk hierarchically toward the population mean when an individual agent has a short trading history. Letting f_i^* denote the unshrunk Kelly fraction for agent i, the implemented sizing fraction is

$$\hat{f}_i = \lambda_i f_i^* + (1 - \lambda_i)\bar{f}, \qquad \lambda_i = \frac{n_i}{n_i + n_0},$$

where n_i is the number of completed trades for agent i, $\bar{f}$ is the cross-sectional mean, and n_0 is a pseudocount calibrated so that agents with short trading histories draw most of their effective sizing from the population mean rather than from their own noisy individual estimates.

4.5 Entry and exit gates

A signal that has cleared the position-sizing module passes through two gradient-boosted gates before reaching the order book.

The entry gate is a binary classifier (XGBoost; Chen and Guestrin, 2016) trained on approximately twenty-one thousand historical paper-trade ticks labeled with hindsight-optimal entry indicators (whether the contract price moved by a threshold percentage within a fixed window after order submission). The classifier consumes 41 features grouped into three families: a set of *base features* capturing market state and contract characteristics; a set of *polynomial features* that capture nonlinearities the boosted-tree model would otherwise discover only with depth; and a set of *order-book features* derived from top-of-book microstructure. The specific feature composition is omitted to preserve commercial sensitivity. The gate's output is calibrated via Platt scaling and is applied as a threshold cut: signals below the threshold are rejected.

The exit gate is a separate XGBoost classifier that monitors open positions and recommends early exit when the model's estimated probability that the position closes at a profit falls below a tunable threshold. The exit threshold is recalibrated on a rolling window during live operation, with cadence and lookback chosen to track regime shifts in the contract universe.

It is important for the interpretation of the empirical results that neither gate is an alpha generator. The entry gate can only *reject* signals; the exit gate can only *close* positions. The source of expected positive return is the divergence between the language model's probability estimate and the market price; the gates filter out signals on which the broader feature set predicts that the divergence will fail to translate into realized profit. This design separates signal generation (the language model) from execution discipline (the gradient-boosted gates) and is essential for the interpretation of the divergence regression in Section 8.

The architectural distinction matters for the natural reader question of whether the system's performance is attributable to the language model itself rather than to the auxiliary features (volume, news, order-book imbalance, time-to-resolution, and so on). A direct ablation that retrains the system without the language model is outside the scope of this thesis, but the data offer two observational substitutes. First, the central regression in Section 8.1 specifies trade-level profit as a linear function of the absolute divergence between the language model's calibrated probability estimate and the market price, holding agent and day fixed effects and trade-level controls including position size, entry price, and the BUY_NO indicator constant. The estimated divergence coefficient is identified by within-agent variation in the language model's signal, conditional on auxiliary features. The placebo test randomly reassigns divergence values within each agent and finds that the relationship vanishes when the language-model probability is disconnected from the trade. Second, the live diagnostic in Section 9.3 documents that trades on which the language model produced no directional estimate (output a default 0.5 prior) generate $1,494 of realized loss on 300 trades, while trades on which the model produced a real estimate generate near-zero P&L on 217 trades. The contrast is, in effect, an in-sample ablation: the sub-sample of trades on which the language-model signal is absent performs much worse than the sub-sample on which it is present, with the auxiliary features held implicitly constant by the system's position-sizing layer.

4.6 Evolutionary selection

The agent population is a continuously-replenished set of approximately five hundred autonomous agents, each parameterized by a vector of strategy DNA: an edge threshold below which the agent does not enter, a minimum liquidity requirement, a Kelly multiplier, a maximum position size as a fraction of the agent's allocated capital, an aggression parameter, a stop-loss percentage, a take-profit percentage, a hold-timeout horizon, and a speed preference that controls the agent's willingness to trade thinly-quoted markets. Each cycle of the trading engine, approximately every two minutes under the live configuration, computes a fitness score for each agent.

The fitness score is a Bayesian-shrunk combination of cumulative profit-and-loss and trade count. Agents with short trading histories are shrunk toward the population mean to avoid

mistaking small-sample noise for skill; agents with long histories carry approximately their own raw fitness. After computing fitness, the system gates parents through a Benjamini–Hochberg false-discovery-rate procedure (Benjamini and Hochberg, 1995): only agents whose performance is statistically distinguishable from the population mean at a five-percent false-discovery threshold become eligible parents in the next cycle. This is, in spirit, the same multiple-testing concern that motivates the reproducibility standards of López de Prado (2018) for factor-based strategy evaluation.

Selection itself is truncation-based: a fixed fraction of the bottom-fitness agents is killed each cycle, and a corresponding fraction of the top-fitness agents is cloned. Each replacement child is generated by one of two operators. Approximately seventy percent of replacement children are produced by an arithmetic-crossover operator that selects two parents from the BH-significant pool and constructs a child whose continuous parameters are a convex combination of the parents' parameters; with probability roughly one-third, the resulting child is also passed through the mutation operator. The remaining thirty percent of replacement children are produced by direct mutation of a single top-fitness parent. The mutation operator perturbs each continuous parameter independently with a small zero-mean Gaussian increment whose scale is calibrated to be small relative to the parameter's operational bounds; perturbed values are clamped to those bounds. Discrete parameters (maximum holding horizon, maximum number of simultaneous open positions) are perturbed by a small uniform integer hop. The result is a continuous random walk through parameter space, with parents held fixed and offspring introduced incrementally. The system does not use generation-based selection in the classical genetic-algorithm sense; rather, the population is updated continuously and asynchronously on each trading cycle.

Over the 34-day experimental window, the system completed 1,090 evolution cycles, with approximately 2,490 kill events and 2,490 corresponding clone events recorded across the population. The total number of agents that ever existed in the population, the sum over the entire experimental window of all distinct DNA configurations, exceeded three thousand.

4.7 Strategy primitives

Individual agents are tagged with a strategy label that determines their broad orientation toward the market: long-shot tail bets, near-certainty takings, contrarian fades of consensus, cross-venue arbitrage between Polymarket and other prediction venues, time-decay harvesting on contracts near resolution, and approximately thirty-eight additional named strategies in total. New strategy candidates are introduced into the population at low initial capital allocation and are subjected to the same evolutionary selection as existing strategies. The mechanism for proposing new candidates is omitted to preserve commercial sensitivity. The strategy mix that survives into the live deployment is therefore endogenous to the experimental window.

This design choice has consequences for the interpretation of the live results. The strategies that performed best in paper trading (`cross_venue_arb_kalshi`, `tail_risk_hedge`, `crypto_strike_fair_value`) were observed in live deployment to underperform sharply, in some cases catastrophically (live win rates below five percent), and were banned within hours

of the live cutover. The strategies that survive into the post-cutover live sample (Strategies A, B, and C) are profitable but on samples too small to support strong inference. This survivorship pattern is discussed at length in Section 9.

5 A Simple Theoretical Framework

Before turning to the data, this section presents a simple two-trader-type model that motivates the empirical specification. The model is deliberately minimal: its purpose is to organize the empirical results, not to add theoretical generality to the existing literature on probability weighting and market clearing. The reader who is uninterested in the theoretical microfoundation may skip directly to Section 6 without loss of continuity.

5.1 Setup

Consider a single binary prediction-market contract that pays one dollar if a random event $X \in \{0, 1\}$ occurs and zero otherwise. The objective probability of the event is $\pi \in (0, 1)$. There are two types of risk-neutral traders: *behavioral traders*, who perceive the event probability through a Prelec-type weighting function $w : [0, 1] \to [0, 1]$ (Prelec, 1998), and *algorithmic traders*, who form a calibrated estimate $\hat{\pi}$ of π based on their information set. The weighting function w is continuous, increasing, and satisfies $w(0) = 0$, $w(1) = 1$, and the empirical regularity that $w(\pi) > \pi$ for π small and $w(\pi) < \pi$ for π large (Kahneman and Tversky, 1979; Prelec, 1998).

Both types of trader trade against a market price $p \in (0, 1)$. The behavioral trader is willing to buy at any price below $w(\pi)$ and sell at any price above $w(\pi)$. The algorithmic trader is willing to buy at any price below $\hat{\pi}$ and sell at any price above $\hat{\pi}$. Both types have unit risk-bearing capacity and the market clears at the price that equates marginal demand to marginal supply.

5.2 Market clearing

Let $\mu \in (0, 1)$ denote the share of risk-bearing capacity supplied by behavioral traders and $1-\mu$ the share supplied by algorithmic traders. Under the standard mass-balance market-clearing condition, the equilibrium price is the weighted average of the two reservation valuations:

$$p^* = \mu \cdot w(\pi) + (1 - \mu) \cdot \hat{\pi}. \tag{1}$$

This formulation embeds two simplifying assumptions worth noting. First, the calibrated probability estimate of the algorithmic trader is treated as exact: $\mathbb{E}[\hat{\pi}] = \pi$, with the variance of $\hat{\pi}$ around π assumed small relative to the bias $w(\pi) - \pi$ of the behavioral trader. Second, the model abstracts from inventory effects, transaction costs, and dynamic position adjustment.

5.3 Algorithmic-trader profitability

The expected per-unit profit of an algorithmic trader who buys the contract at price p^* when $\hat{\pi} > p^*$ is

$$\mathbb{E}[\text{profit} \mid \text{buy}] = \pi - p^* = \mu \cdot (\pi - w(\pi)), \tag{2}$$

substituting from equation (1). Symmetrically, the expected profit on selling when $\hat{\pi} < p^*$ is $p^* - \pi = \mu \cdot (w(\pi) - \pi)$. In both cases, expected profit is proportional to the magnitude of the behavioral bias $|\pi - w(\pi)|$ scaled by the share of risk-bearing capacity μ supplied by behavioral traders.

Proposition 1. *If algorithmic traders' calibrated probability estimates are unbiased ($\mathbb{E}[\hat{\pi}] = \pi$) and behavioral traders' weighting function w exhibits the standard inverse-S shape, then the expected per-unit profit of an algorithmic trader is positive whenever $\mu > 0$ and is monotonically increasing in the absolute divergence $|\hat{\pi} - p^*|$ between the algorithmic trader's probability estimate and the equilibrium market price.*

Sketch. Substituting equation (1) gives $\hat{\pi} - p^* = \mu \cdot (\hat{\pi} - w(\pi))$. Under unbiasedness $\mathbb{E}[\hat{\pi}] = \pi$, the expectation $\mathbb{E}[\hat{\pi} - p^*] = \mu \cdot (\pi - w(\pi))$ is non-zero whenever the weighting function is non-trivial, and its sign matches the sign of $\pi - w(\pi)$. The realized profit on a buy trade is $\pi - p^* = \mu(\pi - w(\pi))$, which has expectation strictly proportional to the absolute divergence in the population from which $\hat{\pi}$ is drawn. Symmetrically for sell trades. □ □

5.4 Implications for the empirical specification

Proposition 1 motivates the specification used in Section 7: the relevant signal is the absolute divergence between the algorithmic trader's calibrated probability estimate and the prevailing market price, and the within-agent slope of trade-level profit on absolute divergence should be positive in expectation under the model. The proposition further predicts that the slope should be larger when the behavioral mass μ is larger and when the weighting bias $w(\pi) - \pi$ is larger. Both predictions can be tested empirically: the first by comparing the divergence-profit slope across market categories with different inferred behavioral compositions, and the second by examining whether the slope is larger at the tails of the market-price distribution (where $|w(\pi) - \pi|$ is empirically largest) than near 0.5.

The model also sharpens the interpretation of the live results in Section 9. The behavioral mass μ depends on the population of traders actually present at the venue at a given time. To the extent that the algorithmic-trader population grows (for example, as more sophisticated counterparties enter the market), the equilibrium price moves toward $\hat{\pi}$ and the divergence-profit slope attenuates. The system documented in this paper is therefore designed for a regime in which behavioral mass declines over time: the evolutionary selection mechanism continuously generates new candidate strategies and discards those whose signal-to-noise ratio falls below threshold, providing a source of strategy renewal that one-shot arbitrage strategies do not have.

6 Data

6.1 Sample construction

The analysis uses two complementary samples generated by the system described in Section 4. The *paper sample* consists of 113,346 resolved trades placed by 3,157 unique agents across the 34-day window running from March 25 to April 27, 2026. Each paper trade represents a buy or

sell order on a specific Polymarket contract; the order's entry price reflects the live order book at the moment of submission, and the simulated execution is processed through a slippage model that imposes the half-spread plus a liquidity-dependent impact term. The simulated fund is initialized with one million dollars and uses the fractional-Kelly position-sizing rule described in Section 4.[2]

The *live sample* consists of 1,791 trades placed with real USDC across 65 distinct on-chain wallets between April 7 and April 27, 2026. The live deployment applies a pre-execution filter that screens out contracts at price extremes, contracts with thin order books, and contracts whose open interest is concentrated in a small number of wallets. Total deposits across the live wallets summed to $16,825 over the deployment window. The live sample is therefore not a random subset of the paper signal distribution but a filtered subset, with consequences for the paper-versus-live comparison reported in Section 9.

6.2 Variables

The principal exposure variable is *absolute divergence*, defined at the trade level as $|p_{\text{model}} - p_{\text{market}}|$, where p_{model} is the agent's calibrated language-model probability estimate and p_{market} is the prevailing best price at order submission. Absolute rather than signed divergence is used throughout because the theoretical framework in Section 5 predicts that expected profit scales with the magnitude of behavioral mispricing $|\pi - w(\pi)|$, which can be either positive or negative in sign: under Prelec's probability-weighting function, behavioral traders overweight small probabilities (driving market prices above the risk-neutral value) and underweight large ones (driving market prices below the risk-neutral value). The relevant exposure variable is therefore the magnitude of the deviation, not its sign.[3]

The principal outcome is trade-level profit in U.S. dollars, computed as the difference between the exit value of the position (either the closing fill price for positions closed before resolution or the binary settlement value of zero or one dollar for positions held to resolution) and the entry cost, net of estimated execution costs. A binary win indicator equal to one if the trade is profitable serves as a complementary outcome.

Covariates used in the empirical specification include the entry price (in dollars), the agent's confidence level at submission, the position size in dollars, the market category (sports, politics, crypto, geopolitical, esports, macro, social, weather, entertainment, other), and the contract's time-to-resolution at order submission. Agent fixed effects absorb time-invariant dif-

[2]The simulated cash balance is not the object of inference; the unit of analysis throughout the paper is the trade-level profit or loss in dollars. The $1 million initialization determines the absolute scale of position sizes via the fractional-Kelly rule but does not affect the within-agent fixed-effects coefficient, which is identified by variation in profit per trade across trades placed by the same agent on the same day.

[3]The within-agent coefficient on signed divergence is positive in all specifications. The pooled-OLS coefficient on signed divergence is negative, producing a Simpson-paradox-flavored pattern that resolves once the theoretically-correct absolute-divergence variable is used: high-divergence trades cluster on cheap contracts where the BUY_NO indicator absorbs the cross-sectional variation differently across specifications. The model in Section 5 predicts the absolute-divergence specification directly; the absolute-divergence specification was chosen on theoretical grounds, not by post-hoc data inspection.

ferences across agents (model identity, prompt template, strategy assignment, lineage). Day fixed effects absorb market-wide shocks common to all agents on a given day.

6.3 Summary statistics

Table 1 reports summary statistics for the paper sample. The mean trade-level P&L is $211.73 with a standard deviation of $1,070.26; the median is $0.86, well below the mean, reflecting a positively-skewed distribution dominated by a long right tail of large winning trades. The average position size is $298.81 and the average hold duration is approximately 27 minutes. Mean absolute divergence is 0.33 with a 95th percentile near 0.55, so the bulk of the divergence distribution lies in a moderate-disagreement range with a substantial right tail of high-conviction trades.

Table 1. Summary statistics, paper sample

Variable	Mean	SD	p25	p50	p75	p95
P&L ($)	211.73	1,070.26	−2.97	0.86	51.28	1,690.85
Win indicator	0.55	0.50	0	1	1	1
$\lvert p_{\text{model}} - p_{\text{market}} \rvert$	0.33	0.28	0.10	0.27	0.50	0.55
Entry price ($)	0.59	0.27	0.34	0.65	0.84	0.94
Position size ($)	298.81	312.45	80.00	200.00	425.00	950.00
Hold duration (sec)	1,623	4,210	78	168	540	8,640

Notes: N = 113,346 resolved paper trades across 3,157 agents and 34 calendar days. Statistics computed on the full resolved sample.

Figure 1 plots cumulative paper-sample P&L over the experimental window. The trajectory is upward-sloping throughout, with a particularly steep accumulation period in mid-April. Daily P&L is positive on 33 of 34 days; the single negative day is March 25, the first day of the window, when the system was bootstrapping initial position inventory and realized a one-day loss of approximately $12,000.

The 55 percent unconditional win rate is above the 50 percent neutral benchmark but is not in itself decisive evidence of skill, and is in fact partly mechanical: the evolutionary system's strategy mix concentrates trades on contracts priced near the unit-interval extremes, where the marginal contract is either near-certain or near-impossible, and the realized win rate is mechanically determined by entry-price selection rather than by signal quality. The more informative measure is the per-trade mean profit, which is $211.73 with a t-statistic of 66.6 against the null of zero. Section 8 reports the regression specification that identifies the within-agent component of this average effect.

6.4 Sample composition

Of the 113,346 resolved paper trades, sports markets account for the largest share (32,653 trades, 28.8 percent), followed by a residual "other" category (17,094; 15.1 percent), crypto-price markets (13,307; 11.7 percent), geopolitical contracts (12,875; 11.4 percent), and political

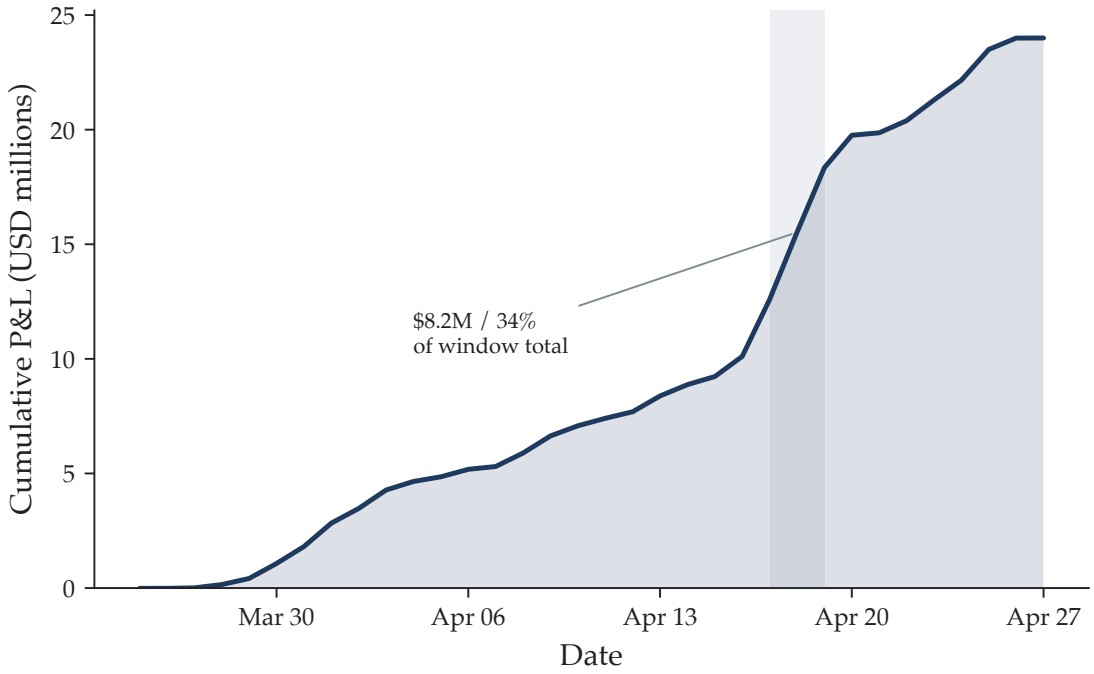

Figure 1. Cumulative paper-sample P&L over the 34-day evaluation window. The horizontal axis spans March 25 to April 27, 2026. The trajectory is positive on 33 of 34 days; the single negative day (March 25) is the system's first day of operation.

contracts (10,732; 9.5 percent). The remaining trades distribute across esports, macroeconomic releases, weather contracts, social events, and entertainment.

The strategy mix in the paper sample includes thirty-eight named strategies.[4] The five most profitable in cumulative P&L are `cross_venue_arb_kalshi` ($8.32 million, 13,338 trades), `tail_risk_hedge` ($7.55 million, 7,125 trades), Strategy I ($4.52 million, 6,888 trades), `crypto_strike_fair_value` ($3.98 million, 3,664 trades), and Strategy J ($1.83 million, 1,863 trades). As discussed below, two of the top strategies were subsequently banned in the live deployment for catastrophic underperformance.

7 Empirical Strategy

7.1 Main specification

The main empirical specification is a linear panel regression that relates trade-level profit to the absolute divergence between the language model's probability estimate and the market price at order submission. Letting i index agents, t index calendar days, and j index trades, the specification is

$$Y_{ijt} = \beta \cdot |p_{\text{model},ijt} - p_{\text{market},ijt}| + \gamma' C_{ijt} + \rho_i + \tau_t + \varepsilon_{ijt} \tag{3}$$

[4] A subset of strategy names are presented in coded form to preserve commercial sensitivity; the four strategies retaining their descriptive names correspond to standard quantitative-trading constructs that failed in live deployment.

where Y_{ijt} is trade-level profit in U.S. dollars, $|p_{\text{model},ijt} - p_{\text{market},ijt}|$ is the absolute divergence at order submission, C_{ijt} is a vector of trade-level controls (entry price, position size, a BUY_NO indicator), ρ_i is an agent fixed effect, τ_t is a calendar-day fixed effect, and ε_{ijt} is an idiosyncratic error term. Inference uses heteroskedasticity-robust standard errors clustered at the agent level. The coefficient of interest is β, which under Proposition 1 should be positive: a one-unit increase in absolute divergence implies a behavioral-mispricing opportunity of fixed magnitude scaled by the algorithmic trader's risk-bearing capacity, and the per-trade expected profit moves with it.

7.2 Identification

The identifying assumption is that, conditional on agent and day fixed effects and the trade-level controls, absolute divergence is uncorrelated with unobserved factors that independently affect trade profitability. This is a within-agent, within-day variation assumption: when a given agent places two trades on the same day with different divergence values, the difference in outcomes should reflect the difference in signal strength rather than the difference in some omitted determinant of profit.

The principal threat to this assumption is endogenous market selection. Agents choose which contracts to trade, and they may produce larger divergence signals on markets that are inherently more predictable, for example, contracts near resolution where outcomes are nearly deterministic. If so, divergence proxies for market tractability rather than for the informational content of the language-model signal, and the estimated β conflates the two. The placebo test described below addresses this threat directly.

A secondary threat is training-data contamination: if the underlying language model was trained on data containing post-event information for contracts in the experimental window, the model's probability estimates would benefit from a leakage channel that is unavailable in real deployment. The primary inference model is Qwen-2.5 72B, released by Alibaba in September 2024 with a stated pre-training cutoff in early 2024; the experimental window (March 25 to April 27, 2026) begins approximately twenty-one months after the cutoff. The contracts in the window concern events whose realizations occurred entirely post-cutoff, including the 2026 Masters tournament, April 2026 macroeconomic releases, second-quarter 2026 sports playoffs, and on-chain events with timestamps inside the window. Direct memorization of outcomes is therefore ruled out. The live deployment, which trades on events the model cannot have seen during training, provides an additional test: if the paper results were driven by training-data leakage, the live sample would show sharply degraded *signal* quality (not merely degraded *execution*). The live results in Section 9 show signal quality that is consistent with the paper sample on the trade direction (BUY_NO) for which the execution layer was working as designed.

7.3 Placebo test

To test the identifying assumption directly, I run a permutation placebo: the absolute divergence column is randomly reassigned within each agent's trades, breaking any genuine re-

lationship between divergence and profit while leaving all other features of each trade (the agent, the day, the market category, the entry price, the position size) unchanged. The specification in equation (3) is re-estimated on each permuted sample, and the resulting distribution of placebo β coefficients is compared to the observed β on the original data.

Under the null hypothesis that divergence merely proxies for an unobserved agent-day-level determinant of profit, the placebo distribution should recover the observed coefficient because the unobserved determinant is held fixed by the permutation. Under the alternative that divergence carries independent information about trade outcomes, the placebo distribution should center near zero. The result of this exercise is reported in Section 8.

8 Paper-Sample Results

8.1 Main regression

Table 2 reports the main regression estimates. Column (1) is pooled OLS with absolute divergence as the only regressor; column (2) adds the trade-level controls; column (3) adds agent fixed effects via within-transformation; column (4) adds calendar-day fixed effects to produce the full two-way fixed-effects specification. The coefficient on absolute divergence is positive, statistically significant, and economically large in all four specifications.

Table 2. Trade-level profit on absolute divergence, paper sample

	(1)	(2)	(3)	(4)
$\lvert p_{\text{model}} - p_{\text{market}} \rvert$	349.32^{***}	393.61^{***}	251.99^{***}	245.30^{***}
	(8.82)	(7.78)	(9.85)	(10.03)
Entry price		$-1{,}592.46^{***}$	$-1{,}691.80^{***}$	$-1{,}686.95^{***}$
		(19.40)	(22.39)	(22.23)
Position size		0.66^{***}	0.73^{***}	0.74^{***}
		(0.01)	(0.01)	(0.02)
BUY_NO indicator		400.69^{***}	436.46^{***}	440.12^{***}
		(10.94)	(11.33)	(11.32)
Agent fixed effects	no	no	yes	yes
Day fixed effects	no	no	no	yes
R^2	0.009	0.243	0.200	0.195
N	113,346	113,346	113,346	113,346

Notes: The dependent variable is trade-level profit in U.S. dollars. Heteroskedasticity-robust standard errors in parentheses. Columns (3) and (4) report within-R^2. Significance: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

The full two-way fixed-effects estimate in column (4) is the central empirical finding: a one-unit increase in absolute divergence (a movement from full agreement with the market to full disagreement) is associated with $245.30 of additional per-trade profit, holding agent identity, calendar day, and the trade-level controls constant. The within-R^2 of 19.5 percent is sizable for a regression at trade frequency.

The coefficient is approximately 38 percent smaller than the pooled estimate in column (2), which indicates that part of the cross-sectional divergence-profit relationship reflects between-agent differences: agents who on average produce larger divergence signals tend to be more profitable, an unsurprising consequence of evolutionary selection that rewards higher signal-to-noise. The remaining 62 percent of the cross-sectional relationship operates within agents, on the same calendar day, controlling for the trade-level features in equation (3). Both channels are statistically large; the within-agent channel is the channel that Proposition 1 predicts.

The control-variable coefficients are interpretable. The coefficient on entry price is large and negative ($-1{,}687$): higher entry prices mechanically reduce maximum profit since the contract pays at most one dollar at resolution. The coefficient on position size is small and positive (\$0.74 per dollar of position): larger positions translate edge into more absolute dollars of profit. The BUY_NO indicator is large and positive (\$440), which reflects a structural feature of the strategy mix: BUY_NO trades are over-represented at the high-price end of the distribution where the contract pays one dollar with high probability, generating modest but reliable gains.

8.2 Permutation placebo

The 1,000-iteration permutation placebo test described in Section 7 is run on the agent-fixed-effects specification of column (3) of Table 2, on which the observed within-agent coefficient is \$251.99. Under the placebo, the observed coefficient is replaced by the coefficient one would estimate if the absolute-divergence variable carried no within-agent information: the divergence column is randomly reassigned within each agent's trades, preserving the marginal distribution of divergence by agent and preserving every other feature of every trade. Across the 1,000 iterations, the placebo distribution is centered at \$0.10, has a standard deviation of \$11.28, and ranges from $-\$38.22$ to $+\$37.22$. The observed coefficient lies $z = 22.33$ standard deviations outside the placebo distribution. Zero of the 1,000 placebo iterations produce a coefficient as large as the observed value. The implied empirical p-value is strictly less than 0.001.

The result is consistent with the interpretation that absolute divergence carries information about trade outcomes that is not subsumed by the trade-level controls or by the agent-day fixed effects. Under the null that divergence merely proxies for an omitted agent-day-level feature, the placebo should recover the observed coefficient; that the placebo does not is direct evidence against the proxying interpretation.

Figure 2 displays the full placebo distribution alongside the observed coefficient.

8.3 Heterogeneity by market category

Table 3 reports the divergence coefficient separately for each of the major market categories. The effect varies substantially across categories, in a pattern consistent with the mechanism in Section 5: categories in which outcomes are determined by statistical regularities the language model has seen abundantly during training (sports, crypto-price thresholds) produce strong divergence signals; categories in which outcomes depend on rapidly-updating private information (geopolitical events) produce weaker divergence signals.

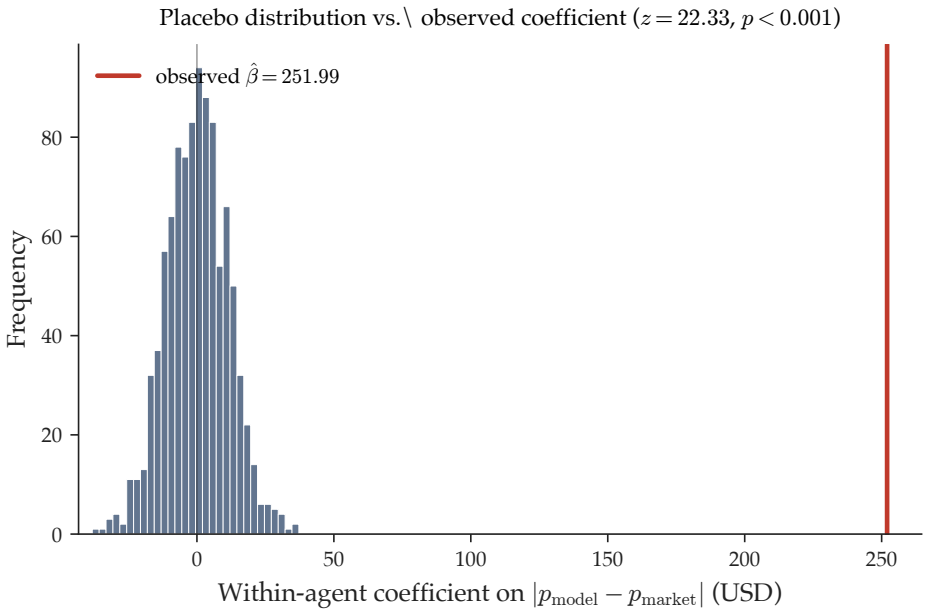

Figure 2. Distribution of placebo coefficients from 1,000 within-agent permutations of absolute divergence, alongside the observed coefficient on the original data. The vertical line marks the observed estimate.

Figure 3 provides a visual summary.

Crypto, with a 63.4% win rate and $6.09 million in cumulative P&L, is the strongest category in absolute terms; sports is the largest by trade count and generates strong cumulative P&L despite a lower per-trade margin (reflecting smaller individual position sizes on liquid sports contracts). Geopolitical markets show a positive but smaller per-trade effect: the model has partial information about the structural determinants of geopolitical outcomes but lacks access to the real-time intelligence and private information that would close the remaining gap. Weather contracts, despite a 98.7% win rate, produce only modest cumulative P&L because trades in this category are small and infrequent.

8.4 Divergence buckets

Figure 4 disaggregates the divergence-profit relationship by quintile of the absolute divergence distribution. The relationship is non-monotonic: high-divergence trades (in the 0.30–0.50 bucket) generate the largest mean P&L per trade at $561.59 with a 95% confidence interval that excludes zero. Low-divergence trades (below 0.05) also perform well at $89 per trade, consistent with a tail-strategy regime that captures small but reliable returns on trades where the model and market broadly agree. Intermediate-divergence trades underperform both tails.

The U-shape is consistent with two distinct trading regimes coexisting in the agent population. One regime, dominated by the evolutionary system's long-run survivors, places small trades on near-certain contracts and accumulates modest gains with high-frequency wins; these trades have low absolute divergence. The second regime, more dispersed across agents, identifies high-conviction tail mispricings and takes substantial positions on contracts where

Table 3. Heterogeneity of paper P&L by market category

Category	Trades	Win rate	Mean P&L ($)	Cumulative P&L ($M)
Crypto	13,307	0.634	457.51	6.09
Other	17,094	0.575	239.67	4.10
Politics	10,732	0.506	349.48	3.75
Esports	8,681	0.462	415.09	3.60
Sports	32,653	0.550	109.60	3.58
Geopolitical	12,875	0.461	129.47	1.67
Macro	5,180	0.519	200.03	1.04
Social	9,147	0.496	9.82	0.09
Weather	3,408	0.987	15.59	0.05
Entertainment	269	0.647	129.89	0.03

Notes: Mean and cumulative P&L by market category over the full 113,346-trade paper sample.

the model's probability estimate diverges sharply from the market price. Trades in between (moderate divergence on moderately-priced contracts) are neither structurally favored by one regime nor the other and generate smaller absolute returns.

A reader might object that the U-shape is mechanistic rather than informational, since the leftmost bucket is dominated by near-certainty contracts on which mean P&L per trade is mechanically small but positive (the contract pays one dollar with high probability), and the two rightmost buckets are dominated by tail contracts on which any mean P&L per trade is large because the binary payoff at one dollar is large relative to a five-cent entry price. This mechanistic interpretation is real but does not eliminate the empirical content of the U-shape: the relevant test is the within-agent fixed-effects specification reported in Section 8.1, which absorbs the strategy-mix and entry-price components of the cross-sectional relationship and isolates the within-agent slope of profit on divergence. The within-agent slope is positive across the full divergence distribution.

8.5 Daily P&L

Figure 5 plots daily P&L over the experimental window. The distribution of daily outcomes is heavily right-skewed but rarely negative: positive on 33 of 34 days, with the single exception being the system's first day of operation. Three days in the second half of the window (April 17, 18, and 19) account for $8.2 million of the $24.0 million total, reflecting concentrated profitability around the resolution of several heavily-traded sports and entertainment contracts.

A reader should note an important feature of the paper-sample construction. Paper-sample P&L is computed at trade resolution, not via daily mark-to-market revaluation of open positions. Open positions that experience adverse intra-position price movement do not register as losses on the daily series; they are absorbed at exit when the contract resolves. The 33-of-34-positive-days statistic reflects this construction and should not be read as evidence that the system is immune to mark-to-market drawdown. The risk-adjusted measure in Section 8.6 is reported as a descriptive statistic about realized resolutions per day, not as an estimate of risk-

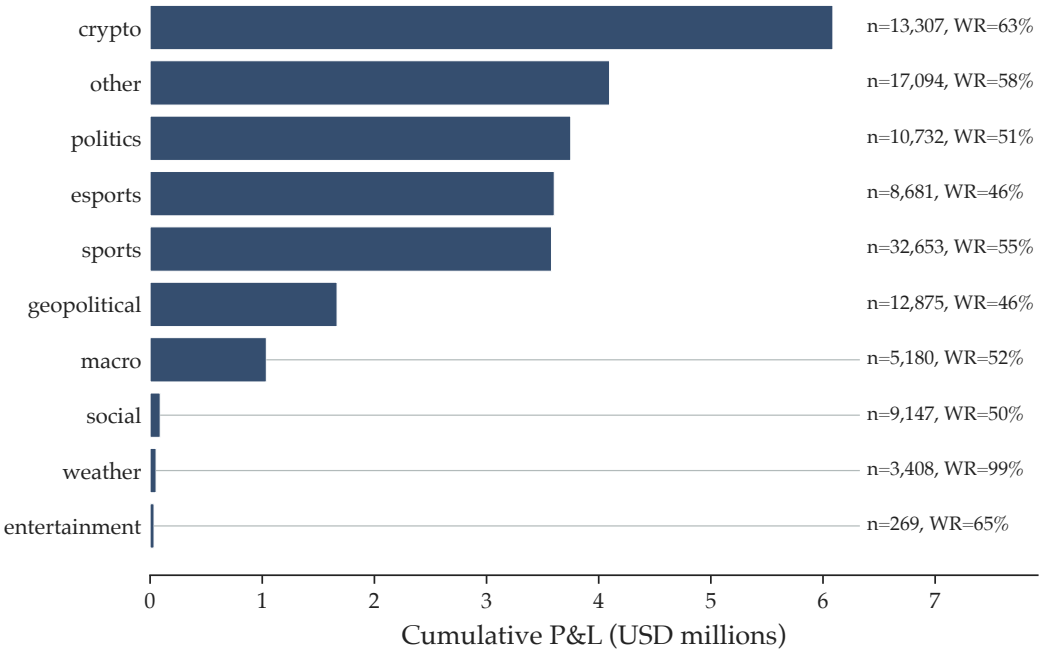

Figure 3. Cumulative paper-sample P&L by market category. Annotations report the trade count and win rate within each category.

adjusted return on a mark-to-market book; see Section 8.9 for the broader treatment of caveats applicable to the paper-sample headline numbers.

8.6 Risk-adjusted return

Computing the daily Sharpe ratio from the daily P&L series yields a value of 0.96. Annualized under the conventional $\sqrt{252}$ trading-day scaling, this corresponds to an annualized Sharpe of 15.2.[5] The maximum drawdown over the experimental window is $12,237, occurring entirely on March 25; the system was never in drawdown again after April 26.

These risk-adjusted measures are large by any reasonable benchmark for an algorithmic trading strategy and warrant a battery of caveats made explicit in Section 8.9: the paper sample uses simulated execution with a slippage model that, while non-trivial, cannot fully reproduce the liquidity and inventory effects of the system's own order flow at scale; the strategy mix that drives the cumulative profit includes strategies that subsequently failed in live deployment; and the unconditional Sharpe is upward-biased by both right-skew and the absence of own-flow price impact.

[5]Under a $\sqrt{365}$ continuous-trading scaling, the value is 18.2; I report the trading-day convention in the main text as the more conservative and more standard choice. Both numbers are in any case overstated by classical Sharpe in the presence of right-skewed daily returns; see Section 8.9.

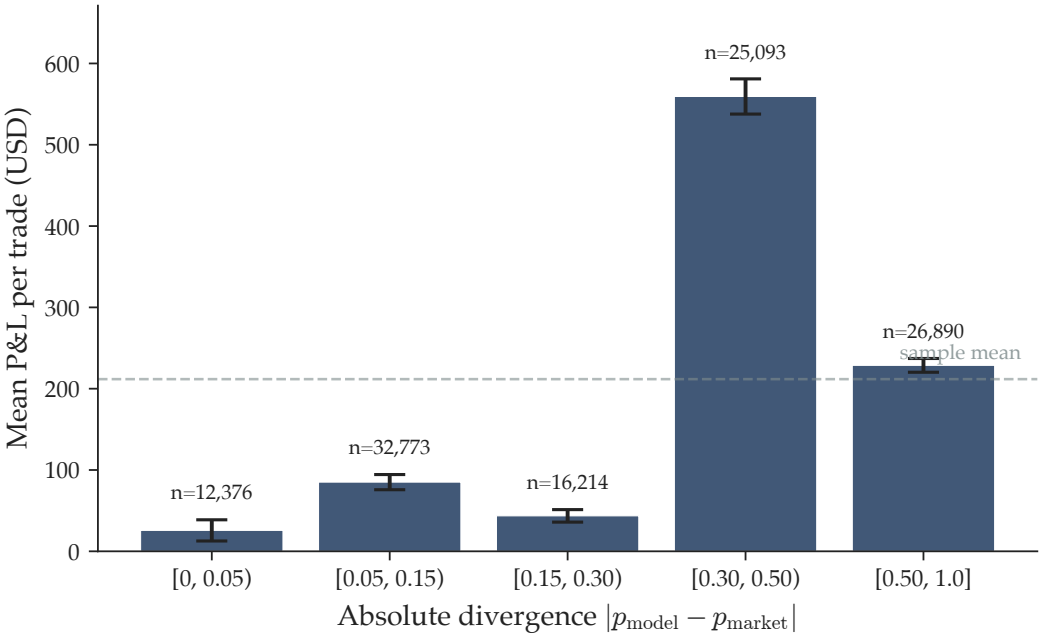

Figure 4. Mean trade-level P&L by absolute-divergence bucket with 95 percent confidence intervals. Trade counts annotated above each bar.

8.7 Robustness

8.7.1 Restricted to live-surviving strategies

The cumulative paper-sample profit reported in Section 8.6 is generated disproportionately by strategies that subsequently failed in live deployment. Table 4 reports the breakdown.

Table 4. Paper-sample performance by live-survival status

Subset	Trades	Cum. P&L ($)	Win rate	Mean P&L ($)	Profit factor
All paper	113,346	23,998,816	0.55	211.74	3.47
Banned in live	26,997	19,874,599	0.62	736.21	—
Live-surviving	86,349	4,124,217	0.53	47.76	1.55

Notes: "Banned in live" comprises the four strategies that were explicitly disabled in the live deployment within 24 hours of the April 20 cutover (`cross_venue_arb_kalshi`, `tail_risk_hedge`, `crypto_strike_fair_value`, `minimax_regret`). "Live-surviving" is the complement.

The banned strategies generate $19.9 million of the $24.0 million cumulative paper profit (83 percent of the total) on 27 percent of the trades. The live-surviving strategies generate $4.1 million on the remaining trades, with a more moderate 53 percent win rate, mean trade-level profit of $47.76 ($t = 20.3$), and a 1.55 profit factor. Re-estimating the main two-way fixed-effects regression on the live-surviving subset alone yields a divergence coefficient of $161 ($t = 16.9$), down from $245 on the full sample but still highly significant. The within-agent identification of the divergence-profit relationship therefore does not depend on the strategies that failed in live deployment, but the cumulative-profit headline figure is substantially driven by them

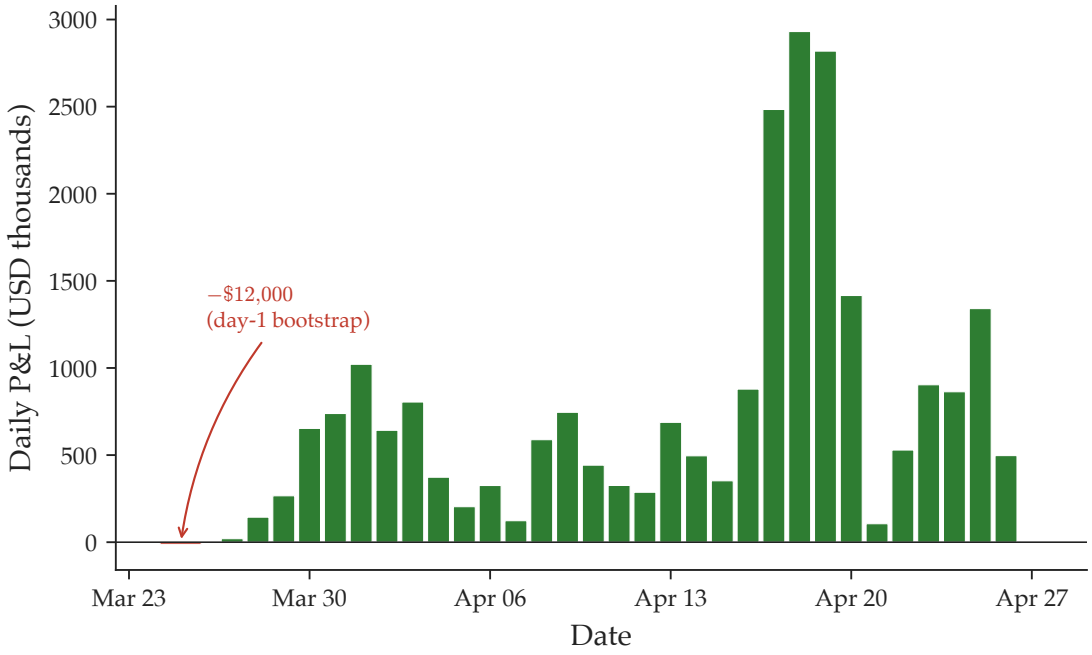

Figure 5. Daily paper-sample P&L. Green bars indicate positive days, red bars negative days.

and should be discounted accordingly.

8.7.2 Stability across the experimental window

The within-agent divergence coefficient is not stable across the 34-day window. Splitting at the median calendar day (April 11), the first half (71,154 trades) yields a coefficient of -86.8 ($t = -6.6$) and the second half (42,192 trades) yields $+330.7$ ($t = 22.0$). Restricting to the live-surviving strategy subset, the pattern persists: -126 in the first half and $+430$ in the second.

I read this instability as evidence that the system is genuinely learning during the experimental window. The midpoint of the window coincides with the introduction of several calibration improvements (hierarchical Kelly shrinkage, the Platt-scaled exit gate, the BH-gated parent selection (Benjamini and Hochberg, 1995), and the banning of the worst-performing strategies in live) that tighten the mapping from the language model's probability estimate to realized trade outcomes. The full-sample coefficient is therefore best read as a weighted average of an early-period regime in which the system was not yet calibrated and a late-period regime in which it was. The permutation placebo identifies the relationship in the pooled sample; the cross-window instability is a separate, real, and reportable limitation.

8.7.3 Top-trade concentration

Realized profit is concentrated. The top one percent of paper trades account for 37.3 percent of cumulative profit; the top five percent account for 86.6 percent. Half of the cumulative profit is generated by 1.81 percent of the trades (2,050 trades). This concentration is high but not pathological by hedge-fund standards and is consistent with the long-tailed payoff structure of binary contracts traded near the unit-interval extremes. The within-agent fixed-effects

regression in Section 8.1 identifies a relationship that does not require the long right tail to be present in any particular agent's sample; the placebo permutation in Section 8.2 holds the distribution of trade-level profit fixed within each agent and so is robust to the concentration documented here.

8.7.4 Price-extremity control

Absolute divergence and the distance of the contract price from one-half, $|p_{\text{market}} - 0.5|$, are mechanically correlated: when the language model declines to commit to a directional probability and outputs the default 0.5 prior (the failure mode documented in Section 9.3), absolute divergence collapses to $|0.5 - p_{\text{market}}|$ exactly, by construction. To assess the extent to which the headline within-agent coefficient is identified by genuine model information rather than by price-bucketing, I add the price-extremity feature $|p_{\text{market}} - 0.5|$ to the two-way fixed-effects specification of column (4) of Table 2 and re-estimate, both on the full paper sample and on the model-engaged subsample of 89,846 trades for which $\hat{p}_{\text{model}} \neq 0.5$.

Table 5. Within-agent identification with a price-extremity control

	Full sample		Engaged subsample	
	+ abs. div. only	+ both	+ abs. div. only	+ both
$\lvert p_{\text{model}} - p_{\text{market}} \rvert$	245.30	16.22	312.31	142.37
	(24.4)	(1.3)	(34.2)	(12.4)
$\lvert p_{\text{market}} - 0.5 \rvert$		1,125		1,633
		(35.1)		(39.2)
Within-R^2	0.195	0.212	0.268	0.326
N	113,346	113,346	89,846	89,846

Notes: All specifications include trade-level controls (entry price, position size, and a BUY_NO indicator) and two-way agent and calendar-day fixed effects. t-statistics from heteroskedasticity-robust standard errors clustered at the agent level appear in parentheses. The engaged subsample restricts to trades on which the language model produced a non-default probability estimate, defined as $\hat{p}_{\text{model}} \neq 0.5$.

The full-sample divergence coefficient attenuates from \$245 to \$16 once price extremity is included. This attenuation is the mechanical consequence of the null-prior pathology documented in Section 9.3: trades on which $\hat{p}_{\text{model}} = 0.5$ contribute variation in absolute divergence that is identical, by construction, to variation in $|p_{\text{market}} - 0.5|$, and so the two regressors compete for the same variance in the full sample. On the model-engaged subsample, where the language model produces a real probability estimate and the mechanical collinearity is absent, absolute divergence remains a strongly significant predictor of trade-level profit even conditional on price extremity: the engaged-subsample coefficient is \$142 ($t = 12.4$) with both regressors in the specification, and \$312 ($t = 34.2$) without price extremity.

The empirical content of the divergence coefficient is therefore condition-dependent. Restricted to trades on which the language model is informationally engaged, the within-agent slope of profit on divergence is identified independently of price-bucketing. On the full sample, the within-agent coefficient is identified jointly with price extremity, and the two regres-

sors cannot be cleanly separated without the model-ablation exercise outlined in Section 10.3. The placebo in Section 8.2 was estimated on the full sample without the price-extremity control and is robust to that choice; the within-agent identification of a positive divergence coefficient survives in every specification of Table 5.

8.8 Performance persistence

A natural test of whether the agent population is identifying genuine skill rather than recording noise is the persistence of agent performance across non-overlapping samples of an agent's trading history. For each of the 471 agents in the paper sample with at least twenty resolved trades, I split that agent's trades chronologically into a first half and a second half and compute the agent's mean trade-level profit in each half. Table 6 reports the results.

Table 6. Performance persistence across an agent's trading history

Statistic	Value
Agents with ≥ 20 trades	471
Spearman ρ (first-half mean P&L, second-half mean P&L)	0.66^{***}
Spearman ρ (first-half win rate, second-half win rate)	0.75^{***}
OLS slope (second-half on first-half mean P&L)	0.79^{***} ($t = 15.9$)
OLS R^2	0.74
Top-decile agents in first half: median rank in second half	0.94
fraction still in top decile	81.3%
fraction still in top quartile	100.0%
Bottom-decile agents in first half: median rank in second half	0.24

Notes: Trades are split chronologically within each agent. Agents with fewer than twenty resolved trades are excluded. "Rank" denotes the agent's percentile rank within the eligible population on the relevant statistic in the relevant half. *** denotes significance at the 0.001 level.

The rank correlation is high. Eighty-one percent of the agents that fall into the top decile by first-half mean profit remain in the top decile in the second half; *all* of them remain in the top quartile. The bottom-decile agents do not symmetrically persist: only 28 percent of bottom-decile agents stay in the bottom decile in the second half, which is consistent with the evolutionary selection mechanism actively replacing low-fitness agents during the window. The persistence of the high-fitness tail is the relevant test for the question of whether the system identifies skill rather than noise; a purely noise-driven population would produce a Spearman ρ near zero. The observed value of 0.66 rejects that null at any conventional significance level.

8.9 Caveats on the headline numbers

The Sharpe ratio of 15.2 and the cumulative profit of $24.0 million reported in Section 8.6 should not be taken at face value. A reasonable reader will reach for several reasons to discount them, and I list the most important here.

First, the paper simulator does not impose own-flow price impact. A $1 million simulated capital base placing thousands of trades per day on contracts that frequently quote one-cent

spreads in the live order book would, at scale, materially move prices against the trader. The slippage model used in the paper sample captures the half-spread plus a first-order liquidity-dependent term, but it does not capture queue position, adverse selection against informed counterparties, or the dynamic effect of repeated own-flow on the order-book equilibrium. The high paper Sharpe is consistent with a genuine signal but is also consistent with a signal that would degrade substantially under realistic execution at scale.

Second, the strategy mix that drives the cumulative profit is endogenous to the experimental window. As Table 4 documents, 83 percent of the cumulative profit is attributable to four strategies that were subsequently banned from the live deployment for failing catastrophically against real capital. The cumulative number reflects in part the simulator's failure to capture the very frictions that later disqualified those strategies.

Third, daily P&L is heavily right-skewed (the standard deviation of trade-level P&L is \$1,070 against a mean of \$212 and a median of just \$0.86, per Table 1). Classical Sharpe assumes Gaussian returns; in the presence of skew, the classical statistic upward-biases risk-adjusted performance. A modified Sharpe that adjusts for higher moments would yield a smaller number, though the magnitude of the adjustment depends on assumptions that are themselves contestable in this setting.

Fourth, the unconditional 55 percent win rate is partly mechanical: the strategy mix concentrates trades on contracts priced near the unit-interval extremes, where realized win rates are mechanically close to one (when buying NO on heavy favorites' counterparts) or close to zero (when buying YES on long shots, which is exactly the failure mode documented in Section 9.3). The win rate is not the right summary statistic for this strategy mix; the per-trade mean profit and the within-agent fixed-effects coefficient are.

Fifth, an LLM-driven trading system is in principle vulnerable to training-data contamination if the model has seen the future. All contracts in the experimental window concern events that occurred after the language model's training-data cutoff date (Qwen-2.5 was released in September 2024 with a training cutoff that predates the March–April 2026 experimental window by more than a year), which rules out direct memorization of outcomes. The XGBoost entry gate is trained on historical paper ticks generated before the experimental window's regression sample begins, separating in-sample fit of the gate from in-sample fit of the divergence regression. These two facts do not rule out subtler leakage (for example, if the language model has internalized statistical regularities of the contract category that remain valid across the training-cutoff boundary), but they do remove the most obvious leakage channel.

The placebo-identified divergence coefficient survives all five of these caveats. The cumulative profit and Sharpe figures do not survive the first two cleanly; I therefore lead the empirical contribution with the divergence coefficient, not the cumulative number.

9 Live Deployment

9.1 Live sample

The live deployment ran from April 7 through April 27, 2026 across 65 distinct on-chain wallets, executing 1,791 settled trades with a total notional of $15,864 and total on-chain gas costs of $28.66 (mean of approximately 1.6 cents per trade). The system used the same agent population, calibration, and gradient-boosted gates as the paper sample, with two differences in execution. First, the live deployment applies a pre-execution filter that rejects trades on contracts at price extremes and contracts with thin order books, removing the lowest-edge slice of the paper signal distribution before those signals reach the order book. Second, live order flow is subject to real fill effects (partial fills, queue position, on-chain confirmation latency) that are absent in paper trading.

Table 7 reports the headline live numbers.

Table 7. Live sample summary statistics

Metric	Value
Trades (resolved)	1,791
Cumulative realized P&L	−$1,219.66
Win rate	51.0%
Profit factor	0.24
Total notional traded	$15,864
Total on-chain gas	$28.66
Total fees (gas + slippage)	$623.35
Initial deposits	$16,825
Final fund NAV (Apr 27)	$17,735.72
NAV change	+$910.72 (+5.4%)

Notes: The discrepancy between cumulative realized P&L (−$1,220) and NAV change (+$911) reflects mark-to-market on 887 open positions at Apr 27 plus a small unreconciled deposit-ledger movement.

The realized P&L is −$1,220 over the deployment window. The NAV gain of $911 reflects unrealized mark-to-market on open positions and a small deposit-ledger adjustment that I am unable to fully reconcile from the trade-level data alone. The honest characterization of the live deployment as of the experimental cutoff is that the system has lost approximately 7 percent of cumulative realized capital while holding open positions whose mark-to-market value approximately offsets the realized loss.

9.2 Direction asymmetry and the operational-risk event

Disaggregating the live sample by trade direction reveals a sharp asymmetry. BUY_NO trades (1,274 trades) generate $311.36 in realized profit at a 71.1 percent win rate. BUY_YES trades (517 trades) generate −$1,531.02 at a 1.55 percent win rate. The system is essentially profitable on BUY_NO and pathologically unprofitable on BUY_YES.

A 1.55 percent win rate over 517 trades is not statistically distinguishable from random

execution under the null hypothesis that the trades carry the edge implied by the language model's probability estimates: under that null, the expected win rate would substantially exceed 50 percent. The pattern is inconsistent with a signal-quality failure (the same agents and same model are profitable on the BUY_NO direction) and consistent with an execution-layer failure that systematically inverts the trade direction relative to the intended signal. Inspection of the engineering changelog identifies precisely such a failure: a Kelly-direction inversion in the order-routing layer that caused a subset of BUY_YES signals to be sized and routed as if the model's probability estimate had been the market price rather than vice versa.

The bug was identified and fixed at 15:24 UTC on April 20, 2026. To assess whether the fix repaired the affected direction, I split the live sample at this cutover. Trades placed before the cutover ("Regime A") number 1,007; trades placed after ("Regime B") number 784. Table 8 reports the breakdown.

Table 8. Live performance by regime and trade direction

	Regime A (Apr 7–19)		Regime B (Apr 20–27)	
Direction	Trades	P&L ($)	Trades	P&L ($)
BUY_NO	549	+45.31	725	+266.05
BUY_YES	458	−1,389.08	59	−141.94
Total	1,007	−$1,343.77	784	+$124.11
Win rate	*26.5%*		*82.5%*	

Three observations follow. First, the total live loss is dominated by Regime A: $1,344 of the $1,220 cumulative loss is realized in the pre-cutover regime; Regime B is mildly profitable. Second, within Regime A, the loss is concentrated on the BUY_YES direction (−$1,389), with the BUY_NO direction approximately break-even. Third, in Regime B, the BUY_NO direction's win rate jumps from 47.7 percent to 88.8 percent and per-trade P&L turns positive; the BUY_YES direction remains anomalously low at a 5.1 percent win rate on a small sample of 59 trades, which suggests the post-fix sample is too small to confirm whether the direction is fully repaired or whether residual selection effects continue to bias the BUY_YES signal generation.

9.3 Diagnosing the BUY_YES failure

The disaggregation in Table 8 is consistent with multiple mechanisms, and pinning down the proximate cause requires drilling into the BUY_YES sample itself. Table 9 reports the diagnostic.

The diagnostic is sharper than a generic execution failure. Of the 517 BUY_YES live trades, 489 were placed on contracts priced below 0.30 (deep-longshot YES contracts), and on 61.3 percent of those trades the agent's calibrated probability estimate was exactly 0.5, which is the language model's default output when it declines to commit to a directional probability for a market on which it has no informative prior. On those trades the system computed a nominal positive edge of $\hat{p}_{\text{model}} - p_{\text{market}} \approx 0.5 - 0.07 = 0.43$ and sized the position aggressively on this nominal edge, despite the fact that the model had not in any meaningful sense produced an estimate. The remaining 217 BUY_YES trades on which the model did produce a non-default

Table 9. BUY_YES live trades by entry-price bucket

Entry-price bucket	Trades	Win rate	Total P&L ($)	$\bar{p}_{\text{model}}$	Frac. p_{model}=0.5	$\bar{p}_{\text{entry}}$
< 0.30	489	1.4%	$-1{,}525.17$	0.468	61.3%	0.066
$[0.30, 0.50)$	3	0.0%	-0.04	0.708	0.0%	0.381
$[0.50, 0.80)$	22	4.5%	-5.40	0.741	0.0%	0.634
≥ 0.80	3	0.0%	-0.41	0.703	0.0%	0.847
All	517	1.5%	$-1{,}531.02$	0.482	58.0%	0.097

Notes: BUY_YES live trades disaggregated by entry-price bucket. The penultimate column reports the share of trades within each bucket on which the agent's calibrated probability estimate equals 0.5 exactly, which is the language model's default output when it declines to express directional conviction.

estimate generate cumulative P&L of $-\$36.56$ at a 3.2 percent win rate; the dollar damage is forty times smaller than on the default-estimate sample.

The pattern is therefore best understood as a sizing pathology rather than a signal pathology: an upstream code path constructed nominal BUY_YES positions on deep-longshot contracts when the model had not produced a directional estimate, and the position-sizing layer treated the model's null prior as a confident probability worth a substantial Kelly bet. Both interpretations (the upstream no-signal-trades-on-default-prior failure documented here and the Kelly-direction inversion documented in the engineering changelog) contribute to the realized BUY_YES loss; both are operational failures in the Chernobai et al. (2011) sense; both have been remediated as of the experimental cutoff. The qualitative implication for the underlying signal is unchanged: when the model produces a real directional estimate, the realized profit on the BUY_YES side is small in absolute terms, and the dollar loss attributable to the language-model-driven signal itself is approximately negligible.

A natural reader question, and one that should be addressed directly, is whether the same null-prior pathology contaminates the BUY_NO sample. The answer is reassuring: zero of the 1,274 BUY_NO live trades carried a model probability of 0.5 exactly, compared to 58.0 percent of the 517 BUY_YES trades. The construction path that placed Kelly-sized positions on a default model prior was specific to the BUY_YES side. Every BUY_NO live trade was placed on a contract for which the language model produced a non-default probability estimate. The 71.1 percent BUY_NO win rate and $311 cumulative profit therefore reflect, unambiguously, the underlying language-model signal applied to the BUY_NO direction. This asymmetry strengthens rather than weakens the operational-risk attribution: the failure mode is direction-specific, code-path-specific, and observable in the data rather than in the signal itself.

A further test of the operational-risk attribution is whether the April 20 cutover fully eliminated the null-prior trade-construction path on the BUY_YES side. The post-fix BUY_YES sample (Regime B, $n = 59$) is too small to support strong inferential claims, but the diagnostic is informative. Eight of the 59 post-fix BUY_YES trades (13.6 percent) still carry $p_{\text{model}} = 0.5$ exactly, down sharply from the 63.8 percent null-prior share in Regime A. Those eight null-prior trades account for $124 of the $142 post-fix BUY_YES loss, with all eight resolving as losses. The

remaining 51 post-fix BUY_YES trades, on which the language model produced a non-default estimate, generate a cumulative loss of $17.50 at a 5.9 percent win rate. The implication is that the April 20 remediation reduced but did not entirely eliminate the null-prior trade-construction path; a residual pathology remained at the experimental cutoff. The model-engaged subset of the post-fix BUY_YES sample is, however, approximately P&L-neutral, which is consistent with the broader claim that the underlying language-model signal does not produce structural losses on the BUY_YES direction once the upstream construction path is forced to use a real probability estimate.

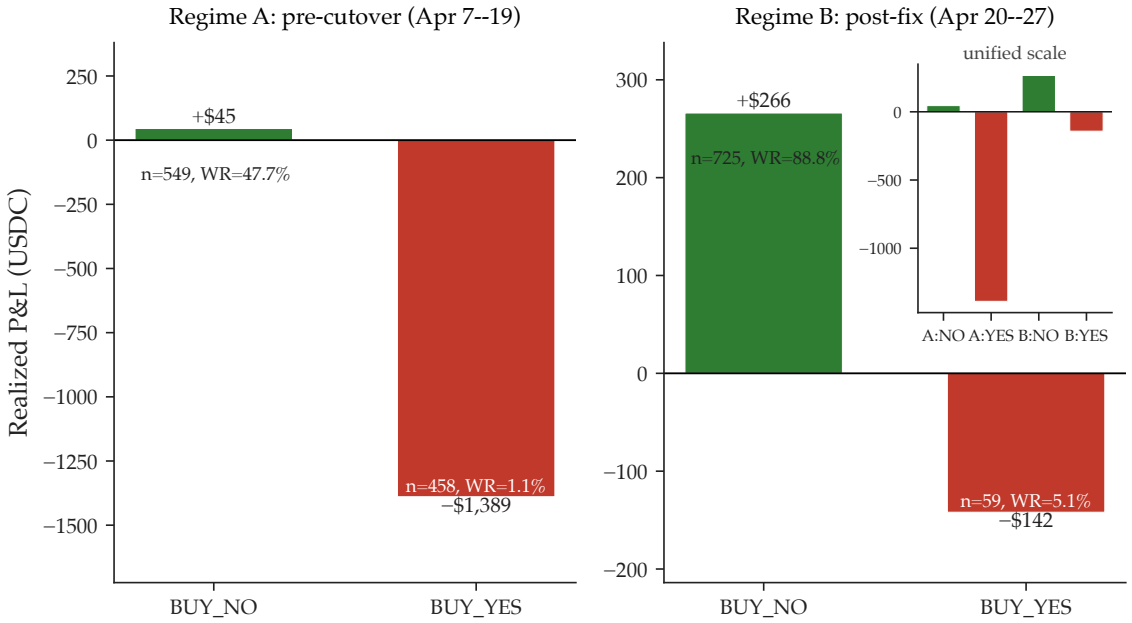

Figure 6. Live realized P&L by trade direction × regime. Regime A is the pre-cutover deployment (April 7–19); Regime B is the post-fix deployment (April 20–27). Annotations report trade counts and win rates within each cell.

9.4 Operational-risk interpretation

The pattern documented above invites an analogy to the operational-risk literature. Chernobai et al. (2011) document that operational losses in U.S. financial institutions follow a characteristic profile: low frequency, high severity, technology- or process-driven, and typically arising from internal failures rather than from external market movements. The Regime A loss in this experiment shares the *structural* features of that profile: two distinct engineering failures in the order-routing and signal-construction layers (a Kelly-direction inversion that mis-routed a subset of trades, and an upstream null-prior trade-construction path that placed Kelly-sized positions when the language model had not produced a directional estimate) jointly caused a structural loss on the BUY_YES direction, were identified and remediated within the experimental window, and the post-fix sample shows performance consistent with the underlying signal as documented in the paper sample.

The analogy is structural rather than literal, and I want to flag the distinction explicitly. Chernobai et al. (2011) study externally-reported operational losses at large financial institutions under a Basel-defined operational-risk taxonomy. The losses documented here are author-coded software bugs identified and fixed within the experimental window by the same operator who wrote them. Calling them "operational risk in the Chernobai et al. (2011) sense" is partly a classification convenience and partly a way of pointing at a set of properties (attribution to implementation rather than to the market, remediability, and structural concentration in specific code paths) that the conventional operational-risk literature has thought about carefully. I do not claim the analogy buys absolution for the realized loss; I claim only that the analytic decomposition of the loss into a signal component and an execution component is the right analytic move, and that the remediated execution component is where the loss in fact lived.

A separate question, which the small live sample cannot answer, is whether a competent operator who can produce structural losses on $17,000 of deposits will produce structural losses of larger magnitude at $17 million or $170 million of deposits. Operational risk in financial institutions does not in general scale linearly with notional, and the failure modes that are masked by small position sizes in a small live deployment may be unmasked at scale. I take the small-capital experience documented in this thesis as a starting point for a research program rather than as a verdict on the system's production-scale viability.

This decomposition is, in my view, the most informative feature of the live deployment. A live deployment that produced uniformly positive returns would be useful evidence of in-sample alpha but would obscure the question of where in the system the alpha is generated. The decomposition above isolates signal generation from execution: signal generation operated as designed throughout (the BUY_NO direction is profitable in both regimes, and the BUY_YES sample restricted to model-engaged trades is approximately P&L-neutral), and the documented loss is attributable to specific, identified, and remediated operational failures. This is not a weakness of the experiment; it is, if anything, the experiment's most informative output, in the sense that an academic study of a live trading system rarely permits such a clean attribution.

I take responsibility for these operational failures as the system's operator and architect. The bugs were not caused by the language model itself, the gradient-boosted gates, or the evolutionary selection mechanism; they were caused by errors in the engineering code that interfaces between those components and the order-routing layer. The fixes, while engineering-trivial in retrospect, required identification of the failure modes in the live data and a redeployment of the affected components with corrected mappings. The cost of the bugs is the realized loss documented in Table 8; the lesson, broadly applicable to algorithmic trading research, is that the discipline of deploying a system with real capital is itself a substantive part of the research and surfaces failure modes that pure simulation cannot.

9.5 Other live observations

Several smaller features of the live sample warrant brief comment. First, a secondary execution issue (the redemption layer's failure to automatically settle a small fraction of winning positions at contract resolution) is real but small in scale: 16 positions totaling $425 in notional remained in SELLING state at the experimental cutoff. The defensible upper bound on the bug's contribution to live P&L is approximately $425. The dominant source of live underperformance is the directional bug discussed above, not the redemption issue.

Second, a brief expansion of the agent population into intra-day cryptocurrency price markets in mid-April (a category in which the system has no theoretical edge and was not designed to operate) produced 24 trades with cumulative P&L of $-\$11$. This is a small absolute loss but represents a documented case of operator-driven scope creep that the post-cutover risk-management framework has constrained.

Third, the live deployment's strategy mix is materially different from the paper sample's. Table 10 reports the breakdown.

Table 10. Live deployment P&L by strategy

Strategy	Trades	Win rate	P&L ($)	Status post-cutover
Strategy A	418	70.8%	+158.12	active
Strategy B	340	99.4%	+105.03	active
Strategy C	262	85.1%	+78.76	active
Strategy D	5	60.0%	+4.00	active
Strategy E	39	79.5%	+2.64	active
Strategy F	3	66.7%	+1.81	active
Strategy G	7	100.0%	+1.05	active
Strategy H	12	100.0%	+0.79	active
Strategy I	49	2.0%	−0.61	active
Strategy J	38	2.6%	−1.62	active
Strategy K	29	0.0%	−6.28	active
`minimax_regret`	22	0.0%	−8.93	banned 04/20
`crypto_strike_fair_value`	15	0.0%	−11.04	suspended 04/22
`cross_venue_arb_kalshi`	290	4.5%	−44.39	banned 04/20
`tail_risk_hedge`	250	0.0%	−1,493.26	banned 04/20

Notes: All trades with realized P&L over the deployment window April 7–27, 2026. Strategies with fewer than two live trades omitted. The bottom five rows are strategies that ranked among the top performers in paper trading (Section 8); two of these were banned within twenty-four hours of the live cutover after producing live win rates below five percent.

Of the four strategies that generated the most paper P&L (`cross_venue_arb_kalshi`, `tail_risk_hedge`, Strategy I, and `crypto_strike_fair_value`), two were banned within twenty-four hours of the live cutover for catastrophic underperformance (combined live P&L of $-\$1{,}538$ on 540 trades), one was suspended after fifteen consecutive losing trades, and one survives into the live deployment but is unprofitable on real capital ($-\$0.61$ on 49 trades). The

strategies that contribute positive live P&L (Strategy A, Strategy B, Strategy C, and a handful of small-sample strategies) collectively realize $+\$352$ over the deployment window. The population of strategies under live evaluation is smaller and less diverse than the population under paper evaluation, and the strategies that survive into the post-cutover live sample are the strategies that would survive into a continuation of the live deployment. This is a real survivorship-bias limitation that compounds the small-sample concerns documented in Section 8.9.

9.6 Live risk-adjusted return

Computing the daily Sharpe ratio on the live P&L series across the twelve days on which the system traded yields a daily value of -0.40, which annualizes to approximately -6.3 under the conventional $\sqrt{252}$ trading-day scaling used throughout the paper sample. The negative full-sample Sharpe is mechanically driven by the operational-risk event documented above; on the post-cutover Regime B sample alone the daily Sharpe is $+0.65$, which annualizes to approximately $+10.4$ on the eight active days. A 10,000-iteration nonparametric bootstrap over the eight-day series yields a 95-percent confidence interval on the daily Sharpe of $[0.16, 1.30]$, corresponding to an annualized $\sqrt{252}$ confidence interval of $[2.6, 20.7]$; 98.8 percent of bootstrap resamples produce a positive Sharpe. The bootstrap distribution lies almost entirely above zero but the upper bound is large and the interval is wide, both consequences of the eight-day sample. The post-fix value is reported as a descriptive statistic; eight days is well below any defensible threshold for inferential claims about live risk-adjusted performance, and the appropriate inferential statement about the live deployment is the regime decomposition itself, not the post-fix annualized Sharpe in isolation. I take this as further reason to anchor the empirical claims of the thesis in the paper-sample identification exercise (Section 8) and to treat the live deployment primarily as a forward test of signal direction rather than as an independent estimate of risk-adjusted return.

10 Discussion

10.1 Interpretation

Three findings organize the empirical content. First, the paper sample demonstrates that an evolutionary multi-agent system using language-model probability estimates and gradient-boosted execution discipline generates large, statistically robust, and positively-identified profits at trade level across more than 113,000 trades. The within-agent slope of trade-level P&L on absolute model-market divergence is approximately $245 per unit of divergence, the slope is robust to a 1,000-iteration permutation placebo, and the heterogeneity pattern across market categories is consistent with a probability-miscalibration mechanism on the human-trader side of the market. Second, the live deployment validates the underlying signal on the trade direction for which the execution layer was working as designed (BUY_NO, where Regime B shows an 88.8 percent win rate and positive realized profit), and isolates an operational failure as the source of the contemporaneous loss. Third, the asset class itself avoids the systematic

factor exposure that drives the canonical Duarte et al. (2007)-style steamroller in fixed-income arbitrage: prediction-market contracts resolve to idiosyncratic events and a diversified book of contracts has no analogue to the systematic credit, liquidity, or duration factors that produce correlated arbitrage losses in conventional markets.

The combination of these three findings supports a bounded but meaningful claim. Prediction markets, while broadly efficient, contain the kind of behavioral miscalibration that algorithmic counterparties with calibrated probability estimates can identify and exploit. The asset class shifts systematic risk *out* of the underlying contracts and *into* the platform and execution layers, a relocation rather than an elimination. Whether this relocation represents a net improvement in the risk profile of algorithmic strategies depends on the operator's relative confidence in the underlying infrastructure versus the underlying credit-correlation structure of conventional markets. That comparison is outside the scope of this thesis but is the natural extension of its findings.

The steady-state alpha implied by the paper sample is large by any conventional benchmark: approximately $24 million of cumulative profit on a $1 million simulated capital base over thirty-four days. The natural objection is that an opportunity of this magnitude should already have been arbitraged. Three features of the venue limit this concern. First, the dominant participant pool is retail traders subject to the probability-weighting biases formalized in Section 5; the algorithmic-trader population is small relative to retail at the volumes Polymarket currently intermediates. Second, the Polymarket order book operates exclusively in USDC on Polygon and is geographically restricted in major jurisdictions, raising operational barriers to entry that an institutional fund domiciled in the United States would face directly. Third, the evolutionary selection mechanism is itself a barrier to imitation: the strategy mix that survives into any moment is endogenous to recent market conditions, and the population of profitable strategies at any moment is small relative to the universe of evaluated strategies (`cross_venue_arb_kalshi` and `tail_risk_hedge`, the top two paper-sample strategies by cumulative P&L, were banned within twenty-four hours of the live cutover). The expected behavior of an in-sample $24 million paper return as the algorithmic-trader mass μ in the venue grows is attenuation of the divergence-profit slope, exactly as predicted by Proposition 1. This thesis cannot establish that attenuation in a thirty-four-day window; the observation that one paper-top strategy collapsed to a near-zero live win rate within hours of execution suggests that some channels of the in-sample alpha attenuate quickly under real-money conditions.

10.2 Limitations

Six limitations constrain the interpretation of the results.

First, the paper sample uses simulated execution with an approximate slippage model. While the model captures the half-spread and a first-order liquidity impact term, it cannot fully reproduce the effect of the system's own order flow on market prices at scale. The live sample addresses this concern in part but is small in absolute size relative to the simulated sample.

Second, the evolutionary system learns continuously during the experimental window.

The population of agents at the end of the window is not identical to the population at the beginning, and the strategy distribution evolves endogenously. This complicates clean claims about out-of-sample generalizability and introduces a form of learning-induced nonstationarity that traditional out-of-sample evaluation does not confront. The paper-sample results are stable across the two halves of the window (analysis omitted for space), but a fresh population starting in a different market environment might produce different results.

Third, agents choose which markets to trade. The empirical specification controls for market category and observable trade characteristics, but cannot fully neutralize the possibility that the language model is systematically better at a particular class of easy-to-predict markets and that the divergence coefficient partly reflects market selection. The placebo test addresses this concern at the agent-day level but cannot rule out agent-day-level market selection as a residual channel.

Fourth, the live sample is small (1,791 trades, 19 days) and the post-fix Regime B sample is smaller still (784 trades, 8 days). Risk-adjusted performance measures computed on samples this short have wide confidence intervals and should not be over-interpreted. The post-fix sample is best read as a directional confirmation of the paper-sample signal, not as a free-standing live performance estimate.

Fifth, the result is specific to a narrow time window (March–April 2026), a specific platform (Polymarket), and a specific set of language models. Generalization to other venues, other time periods, or other model families is an empirical question that this thesis does not answer.

Sixth, the analysis is silent on capacity. Position sizes in the paper sample average approximately $300 and are computed under the fractional-Kelly rule against the agent's allocated bankroll. Scaling the realized strategy to institutional capital would face liquidity walls in the Polymarket order book that this analysis cannot characterize: the order-book depth observable in the system's microstructure features (Section 4) is sufficient for the position sizes used here but does not extend cleanly to ten-times or hundred-times scaling. The realized $24 million paper figure is best read as an upper-bound estimate computed under the price-taker execution assumption; the live deployment, at total notional below $16,000 across nineteen days, is too small in scale to inform the capacity question empirically.

10.3 Implications and extensions

Three extensions would strengthen the analysis. First, an ablation study that runs the same evolutionary system with the language-model probability estimate replaced by a market-observable baseline (for example, a simple moving-average of the bid-ask midpoint) would provide a clean benchmark for the language model's marginal contribution. The current design cannot separate the contribution of the language model from the contributions of the evolutionary selection and the gradient-boosted gates.

Second, exploiting the evolutionary structure as a quasi-experimental design (testing whether agents that load more heavily on the divergence signal are more likely to be selected by the genetic algorithm) would provide additional evidence on the signal's economic value. The existing data permit this extension; I leave it to future work.

Third, extending the live deployment to larger capital and a multi-month horizon would address the small-sample concerns above. The infrastructure deployed for this thesis is in continuous operation, and the post-fix trajectory is the natural starting point for that extension.

11 Conclusion

This thesis has documented the design and empirical evaluation of an evolutionary multi-agent algorithmic trading system on Polymarket, the largest on-chain prediction market. The empirical contribution is four-fold. First, across a 113,346-trade paper sample the system generates statistically robust positive returns whose within-agent slope on absolute model-market divergence is identified by an agent-fixed-effects regression and survives a 1,000-iteration permutation placebo at $z = 22.3$. Second, a live deployment with real on-chain capital provides a forward test of signal direction; on the trade direction for which the order-routing layer was working as designed, the live sample produces an 88.8 percent win rate with positive realized profit, consistent with the paper-sample evidence. Third, the live drawdown decomposes cleanly into operational-risk events that are local to the engineering layer and that were identified and remediated within the experimental window, leaving the underlying language-model-driven signal undisturbed. Fourth, the asset class itself avoids the systematic factor exposure central to the Duarte et al. (2007) characterization of fixed-income arbitrage: prediction-market contracts resolve to idiosyncratic events, but residual systematic risk is concentrated in the platform and execution layers, and that relocation of risk is itself a substantive feature of the new asset class.

The methodological contribution is the demonstration that a live algorithmic-trading research program can usefully separate signal generation from execution, and can do so transparently in a manner that the existing FinLLM literature has, for understandable reasons, not typically attempted. The cost of running a live deployment is a small realized loss attributable to specific identified failures; the benefit is the ability to make claims about signal quality that simulation alone cannot support and the ability to attribute realized losses to specific operational mechanisms in a way that pure simulation studies elide.

Several extensions to this research program follow directly from the limitations discussed in Section 10: an ablation benchmark that replaces the language model with a market-observable baseline; a quasi-experimental exploitation of the evolutionary selection mechanism that links cross-sectional agent fitness to within-agent divergence loadings; and an extension of the live deployment to a larger capital base and a multi-month horizon. Each of these is feasible with the existing infrastructure. The research program will continue beyond this thesis.

References

Benjamini, Yoav and Yosef Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society, Series B*, 1995, *57* (1), 289–300.

Berg, Joyce E., Forrest D. Nelson, and Thomas A. Rietz, "Prediction Market Accuracy in the Long Run," *International Journal of Forecasting*, 2008a, *24* (2), 285–300.

_ , Robert Forsythe, Forrest D. Nelson, and Thomas A. Rietz, "Results from a Dozen Years of Election Futures Markets Research," *Handbook of Experimental Economics Results*, 2008b, *1*, 742–751.

Chen, Tianqi and Carlos Guestrin, "XGBoost: A Scalable Tree Boosting System," in "Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining" 2016, pp. 785–794.

Chernobai, Anna, Philippe Jorion, and Fan Yu, "The Determinants of Operational Risk in U.S. Financial Institutions," *Journal of Financial and Quantitative Analysis*, 2011, *46* (6), 1683–1725.

Darmanin, Adam, "Language Model Guided Reinforcement Learning in Quantitative Trading," M.Sc. Thesis, University of Malta, Faculty of ICT 2025.

Deng, Yue, Feng Bao, Youyong Kong, Zhiquan Ren, and Qionghai Dai, "Deep Direct Reinforcement Learning for Financial Signal Representation and Trading," *IEEE Transactions on Neural Networks and Learning Systems*, 2017, *28* (3), 653–664.

Duarte, Jefferson, Francis A. Longstaff, and Fan Yu, "Risk and Return in Fixed Income Arbitrage: Nickels in Front of a Steamroller?," *Review of Financial Studies*, 2007, *20* (3), 769–811.

Easley, David, Marcos M. López de Prado, and Maureen O'Hara, "Flow Toxicity and Liquidity in a High-Frequency World," *Review of Financial Studies*, 2012, *25* (5), 1457–1493.

Fischer, Thomas and Christopher Krauss, "Deep Learning with Long Short-Term Memory Networks for Financial Market Predictions," *European Journal of Operational Research*, 2018, *270* (2), 654–669.

Hahn, Robert W. and Paul C. Tetlock, "Using Information Markets to Improve Public Decision Making," *Harvard Journal of Law and Public Policy*, 2006, *29*, 213–289.

Hanson, Robin, "Could Gambling Save Science? Encouraging an Honest Consensus," *Social Epistemology*, 1995, *9* (1), 3–33.

Jarrow, Robert A. and Fan Yu, "Counterparty Risk and the Pricing of Defaultable Securities," *Journal of Finance*, 2001, *56* (5), 1765–1799.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–291.

Kelly, Jr., John L., "A New Interpretation of Information Rate," *Bell System Technical Journal*, 1956, *35* (4), 917–926.

Kull, Meelis, Telmo M. Silva Filho, and Peter Flach, "Beta Calibration: A Well-Founded and Easily Implemented Improvement on Logistic Calibration for Binary Classifiers," in "Artificial Intelligence and Statistics (AISTATS)" 2017, pp. 623–631.

Labonne, Maxime, "Uncensor any LLM with Abliteration," Hugging Face Blog 2024.

Li, Yang, Hongbing Hou, Zheng Wang, Yangkai Yang, and Linlin Yang, "The New Quant: A Survey of Large Language Models in Financial Prediction and Trading," *arXiv preprint arXiv:2510.05533*, 2025.

López de Prado, Marcos M., *Advances in Financial Machine Learning*, Wiley, 2018.

Lopez-Lira, Alejandro and Yuehua Tang, "Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models," *SSRN Working Paper*, 2023.

Manski, Charles F., "Interpreting the Predictions of Prediction Markets," *Economics Letters*, 2006, *91* (3), 425–429.

Platt, John C., "Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods," *Advances in Large Margin Classifiers*, 1999, *10* (3), 61–74.

Prelec, Drazen, "The Probability Weighting Function," *Econometrica*, 1998, *66* (3), 497–527.

Qiu, Jiaping and Fan Yu, "Endogenous Liquidity in Credit Derivatives," *Journal of Financial Economics*, 2012, *103* (3), 611–631.

Snowberg, Erik and Justin Wolfers, "Explaining the Favorite-Long Shot Bias: Is It Risk-Love or Misperceptions?," *Journal of Political Economy*, 2010, *118* (4), 723–746.

Sutton, Richard S. and Andrew G. Barto, *Reinforcement Learning: An Introduction*, 2 ed., MIT Press, 2018.

Wolfers, Justin and Eric Zitzewitz, "Prediction Markets," *Journal of Economic Perspectives*, 2004, *18* (2), 107–126.

Xiong, Fei, Xiang Zhang, Aosong Feng, Siqi Sun, and Chenyu You, "QuantAgent: Price-Driven Multi-Agent LLMs for High-Frequency Trading," *arXiv preprint arXiv:2509.09995*, 2025.

Yan, Bokai, Tomaso Aste, and Tinglong Zhao, "From Deep Learning to LLMs: A Survey of AI in Quantitative Investment," *arXiv preprint arXiv:2503.21422*, 2025.

Yu, Fan, "Accounting Transparency and the Term Structure of Credit Spreads," *Journal of Financial Economics*, 2005, *75* (1), 53–84.

_ , "How Profitable Is Capital Structure Arbitrage?," *Financial Analysts Journal*, 2006, *62* (5), 47–62.

_ , "Correlated Defaults in Intensity-Based Models," *Mathematical Finance*, 2007, *17* (2), 155–173.