POSTR: POSITION TRANSITION REASONING FOR MULTI-AGENT FINANCIAL DECISIONS

Anonymous authors
Paper under double-blind review

ABSTRACT

Multi-agent financial trading systems typically improve decision quality by assigning specialized agents to analyze heterogeneous information sources, such as market signals, historical trends, and fundamental factors. However, their final decisions are often compressed into coarse-grained actions, such as buy, hold, and sell, resulting in a mismatch between rich analytical reasoning and simplified trading execution. When an aggregator receives only such action labels, distinct target-position transitions and their implied market assessments are treated as the same action, a loss we term *Decision Semantic Collapse*. To address this limitation, we propose **PosTR**, a **Pos**ition **T**ransition **R**easoning framework that reformulates multi-agent trading from action prediction into position-transition reasoning. Instead of directly aggregating discrete trading actions, PosTR introduces an ordered ***position-state space***, where each agent provides a target position, its transition from the previous target position, and a distribution over candidate target positions. This formulation preserves the target-position transition and its implied market assessment while representing uncertainty over the current target position. Furthermore, PosTR develops a ***transition-aware credibility mechanism*** that dynamically evaluates agent reliability. The learned credibility scores reweight trader distributions during consensus formation, while settled branch performance controls the strength of the credibility-based correction. Experiments with role-specialized agents on eight equities across diverse industries show stronger net and risk-adjusted performance than endpoint-only, static-weight, and history-aware baselines. These results highlight that effective multi-agent trading requires reasoning over position-state transitions rather than merely aggregating discrete trading actions.

1 INTRODUCTION

Recent LLM-based multi-agent trading systems organize agents into specialized analytical roles that provide complementary market evidence (Ding et al., 2026; Dong et al., 2025; Yu et al., 2024; Xiao et al., 2024; Zhang et al., 2024; Li et al., 2025c). Because these analyses must be synthesized into a final position, decision quality depends on how trader decisions are represented and aggregated. Isolated action labels create a representation bottleneck because distinct transitions between ordered position states can collapse into the same action.

Existing systems coordinate these analyses through hierarchical management, deliberation, memory, and outcome feedback (Yu et al., 2025; 2024; Xiao et al., 2024; Li et al., 2024; Xiong et al., 2025; Li et al., 2025c; Tian et al., 2025), yet generally leave the coarse action representation used for aggregation unchanged. Compressing rich analyses into Buy, Hold, or Sell collapses distinct transition paths and their implied market assessments. We call this many-to-one compression of explicit transition structure and implicit market assessment *Decision Semantic Collapse*. Because this information is lost before aggregation, it cannot be recovered from the action label alone. With an action-only aggregation interface, we therefore preserve the ordered transition before estimating credibility or forming consensus.

As illustrated in Figure 1, $S \to F$, $F \to L$, and $S \to L$ may all imply Buy, while $L \to F$ and $L \to S$ may both imply Sell. These transitions differ in path and in the market assessments implied by their ordered states. Our key insight is to treat an action as the consequence of a position-state

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

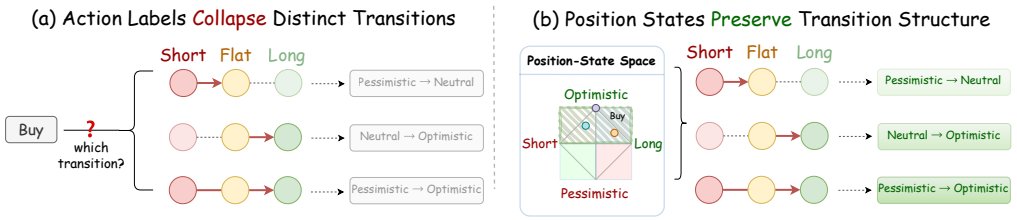

Figure 1: **Action-based versus position-state representation.** (a) Coarse actions collapse distinct position-transition paths. (b) PosTR explicitly represents each path, which implicitly conveys the trader's pessimistic, neutral, or optimistic assessment of market conditions and its relative strength.

transition. The origin and destination explicitly identify the path, while the ordered direction and span provide its implied market assessment. The same space supports comparison and consensus across heterogeneous trader decisions.

Motivated by this insight, we develop **PosTR**, a Position Transition Reasoning framework that addresses *Decision Semantic Collapse*. PosTR represents each trader through its target-position transition and a distribution over candidate positions. The transition preserves the path and its implied market assessment, while the distribution represents uncertainty. PosTR defines aggregation over this shared space by estimating credibility from each structured representation, its peer relations, current market conditions, and settled performance. These space-conditioned weights aggregate the state distributions into a final consensus, whose system-position change produces the executable action. Credibility and branch performance are updated only after the corresponding outcome is observed.

We summarize our contributions below.

- We formulate multi-agent financial decision synthesis as space-conditioned aggregation in a shared **Position-State Space** that preserves position-transition paths and their implied market assessments while supporting system-level consensus.
- We develop a **Transition-Aware Credibility Mechanism** that estimates context-dependent trader reliability and adaptively controls the credibility-based correction.
- We evaluate PosTR on eight equities across diverse industries, demonstrating consistent gains in profitability and risk-adjusted performance against matched baselines and ablations.

2 Related Work

LLM-Based Financial Trading Agents. LLM-based financial trading agents have progressed from individual pipelines to organizational multi-agent systems. Individual agents use memory, multimodal perception, tools, reflection, and outcome feedback to combine heterogeneous evidence (Shinn et al., 2023; Yu et al., 2025; Zhang et al., 2024; Li et al., 2024; Xiong et al., 2025). Related financial-agent pipelines also include reinforcement-learning trading and portfolio platforms (Wang et al., 2021; Sun et al., 2023). Multi-agent frameworks distribute fundamental, technical, sentiment, and risk analysis across specialized roles and coordinate them through management, debate, and structured meetings (Yu et al., 2024; Xiao et al., 2024; Li et al., 2025c; Hong et al., 2024; Wu et al., 2023; Chen et al., 2024b). Other systems address hedging, position-aware management, and fine-grained task decomposition (Li et al., 2025b; Liu & Dang, 2025; Miyazaki et al., 2026), while benchmarks evaluate LLM-based financial decision-making (Li et al., 2025a) and surveys summarize the field's development (Ding et al., 2026; Dong et al., 2025). Despite richer reasoning, these systems generally expose downstream aggregation to a current action or target position rather than a position-transition path and its implied market assessment.

Multi-Agent Decision Coordination. Multi-agent coordination includes proposal selection, deliberation, reliability modeling, belief-based interaction, resilience, reputation, probability pooling, and online expert weighting (Bradley & Terry, 1952; Wang et al., 2022; Jiang et al., 2023; Chen

et al., 2024a; Choi et al., 2026; Li et al., 2023; Du et al., 2023; Chan et al., 2024; Wang et al., 2025; Yi et al., 2025; Mu et al., 2026; Huang et al., 2024; Ren et al., 2025; Xia et al., 2026; Zhang et al., 2025; Genest & Zidek, 1986; Littlestone & Warmuth, 1994; Freund & Schapire, 1997; Hazan et al., 2007; Luo & Schapire, 2015). These methods determine influence from current proposals, interaction, or historical performance, but rarely condition trader reliability jointly on cross-time position transitions and contemporaneous market and peer context.

Structured Representations of Agent Decisions. Structured representations preserve dependencies lost in isolated action labels. Sequence models encode trajectories and long-horizon temporal context (Chen et al., 2021; Janner et al., 2021; Wen et al., 2022), relational methods capture dependencies among heterogeneous agents (Das et al., 2019; Utke et al., 2025), and calibration and evidential learning represent categorical uncertainty (Guo et al., 2017; Kull et al., 2019; Sensoy et al., 2018). These lines address temporal evolution, inter-agent relations, and uncertainty largely in isolation. They leave unresolved how to represent transition paths, preserve their implied market assessments, and condition reliability-aware aggregation on these structures in one space.

3 Problem Formulation

3.1 Sequential Multi-Agent Trading Setting

We consider a sequential trading system with N role-specialized traders. At each decision step t, trader i produces a target position

$$s_{i,t} \in \mathcal{S} = \{L, F, S\}, \tag{1}$$

where L, F, and S denote Long, Flat, and Short, respectively, with $S \prec F \prec L$ representing increasingly optimistic assessments of market conditions. The system aggregates the traders' decisions into an executable position $x_t^* \in [-1, 1]$. Positive, zero, and negative values denote long, flat, and short system positions, respectively.

Trading decisions evolve sequentially rather than in isolation. We therefore retain each trader's previous target position $s_{i,t-1}$ and define its position transition as

$$\tau_{i,t} = (s_{i,t-1}, s_{i,t}). \tag{2}$$

Each trader also provides a distribution over its candidate target states,

$$\mathbf{P}_{i,t} = (p_{i,t}^L, p_{i,t}^F, p_{i,t}^S) \in \Delta^2, \qquad \sum_{a \in \{L,F,S\}} p_{i,t}^a = 1. \tag{3}$$

The reported target position $s_{i,t}$ is the mode of $\mathbf{P}_{i,t}$. The transition captures the path and its implicitly conveyed assessment of market conditions, whereas $\mathbf{P}_{i,t}$ represents uncertainty over the current target position. Here, $s_{i,t}$ denotes the position preferred by an individual trader, whereas x_t^* denotes the system-level position produced after multi-agent aggregation. The system position is determined using only information available at step t, before the next-period outcome is observed.

3.2 Decision Semantic Collapse

Existing multi-agent trading methods commonly express each trader's final decision as a discrete trading action, typically Buy, Hold, or Sell. Under a directional action convention, these labels correspond to upward, zero, and downward changes in the ordered position-state space. The mapping is many-to-one: the labels do not identify the position from which an adjustment starts or the position it is intended to reach. For example, $S \to F$, $F \to L$, and $S \to L$ can all map to Buy, although they represent pessimistic-to-neutral, neutral-to-optimistic, and pessimistic-to-optimistic assessment shifts, respectively. Likewise, $L \to F$ and $L \to S$ can both map to Sell, although they represent optimistic-to-neutral and optimistic-to-pessimistic assessment shifts, respectively. Reducing these transitions to the same action label makes their explicit paths indistinguishable and compresses their associated assessment shifts into the same coarse category. We refer to this coupled representational loss as *Decision Semantic Collapse*.

This limitation motivates us to represent trader decisions by target positions rather than isolated actions. A target position specifies where the trader intends to be, and two consecutive target positions uniquely determine the transition $s_{i,t-1} \to s_{i,t}$. The transition explicitly preserves the decision

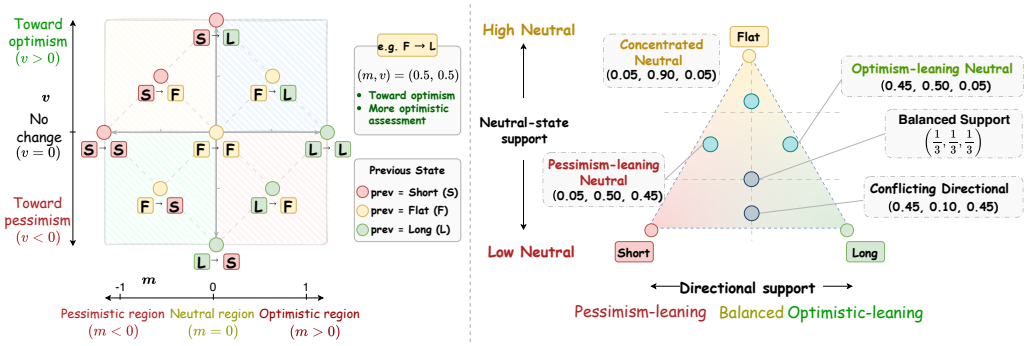

Figure 2: **Position-state transition and distributional representations.** (a) Transition coordinates separate transition location from signed target-position change. (b) State distributions represent uncertainty over candidate target positions.

path. Because the position states are ordered, its direction and span implicitly convey whether the trader's assessment of market conditions becomes more pessimistic or optimistic and the extent of that change.

4 METHOD

PosTR defines a space-conditioned aggregation rule over structured trader representations. Each trader is represented by an ordered target-position transition and a distribution over candidate target positions. The transition explicitly records the path, while its ordered structure implicitly conveys the trader's assessment of market conditions. The distribution represents uncertainty and supplies the consensus payload.

4.1 POSITION-STATE SPACE

Building on the target-position formulation above, we encode Short, Flat, and Long as

$$\psi(S) = -1, \qquad \psi(F) = 0, \qquad \psi(L) = 1. \tag{4}$$

For simplicity, we use $s_{i,t}$ to denote the corresponding encoded value in the following. Each trader also provides $\mathbf{P}_{i,t} = (p^L_{i,t}, p^F_{i,t}, p^S_{i,t}) \in \Delta^2$, ordered as (Long, Flat, Short). The ordering $S \prec F \prec L$ represents increasingly optimistic assessments of market conditions, while $\mathbf{P}_{i,t}$ represents uncertainty over the current target position.

Position-State Transition. For trader i, the ordered pair $\tau_{i,t} = (s_{i,t-1}, s_{i,t})$ records its position transition across two consecutive decision steps. Although the pair uniquely identifies the transition, it does not directly organize the relations among different paths. For example, $L \to F$ and $S \to F$ share the same destination but move in opposite directions in the ordered position-state space, whereas an endpoint-based comparison treats them as the same Flat decision.

We therefore reparameterize each transition by its location and signed target-position change:

$$m_{i,t} = \frac{s_{i,t-1} + s_{i,t}}{2}, \qquad v_{i,t} = \frac{s_{i,t} - s_{i,t-1}}{2}. \tag{5}$$

The coordinate $m_{i,t}$ identifies where the transition occurs, while the sign and magnitude of $v_{i,t}$ describe the direction and extent of the target-position change. Positive and negative $v_{i,t}$ correspond to movement toward more optimistic and more pessimistic assessments of market conditions, respectively.

The two coordinates separate what the endpoint pair leaves entangled. $L \to F$ and $S \to F$ have opposite $v_{i,t}$, so $v_{i,t}$ separates revisions that arrive at the same position from opposite sides. Conversely, $S \to F$ and $F \to L$ have the same positive $v_{i,t}$ but different $m_{i,t}$, so $m_{i,t}$ distinguishes their

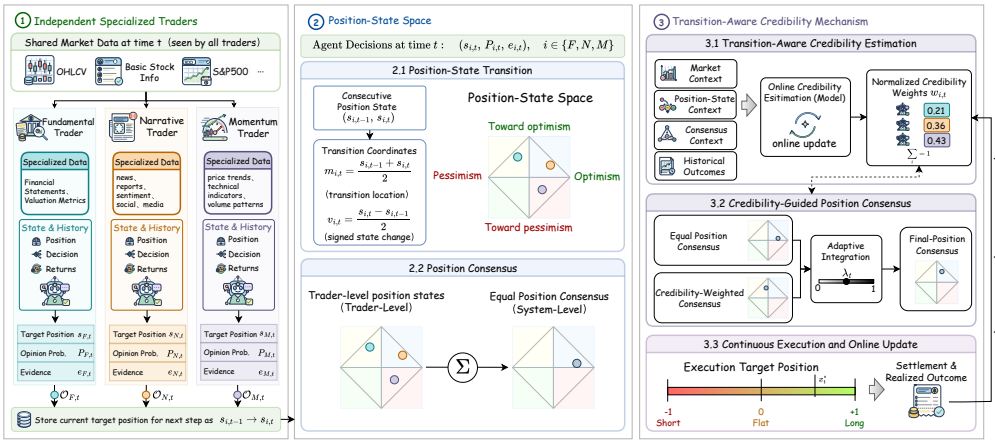

Figure 3: **Overall framework of PosTR.** Each validated trader output $\mathcal{O}_{i,t} = (s_{i,t}, \mathbf{P}_{i,t}, e_{i,t})$ comprises a target, state distribution, and evidence. Transitions and distributions guide aggregation; settled outcomes update credibility. Evidence $e_{i,t}$ is retained but not directly used in aggregation.

transition regions while preserving their common movement toward a more optimistic assessment. The mapping is invertible, with $s_{i,t-1} = m_{i,t} - v_{i,t}$ and $s_{i,t} = m_{i,t} + v_{i,t}$, so it adds no information to the endpoint pair; it reorganizes the pair so that direction and location can be compared across traders. Figure 2(a) illustrates these paths among Short, Flat, and Long.

The transition coordinates make the path explicit and convey its implied assessment direction, but they do not express how firmly the trader holds its current target. The state distribution supplies this uncertainty: two traders with the same transition may concentrate their probability on the reported target or split it toward an adjacent position (Figure 2(b)). Together, they define the structured trader representation

$$\zeta_{i,t} = (m_{i,t}, v_{i,t}, \mathbf{P}_{i,t}). \tag{6}$$

Position Consensus. The shared space allows heterogeneous trader representations $\zeta_{i,t}$ to be compared without reducing them to Buy, Hold, or Sell actions. Transition coordinates, together with relations to peer representations and the equal consensus, determine how each trader is evaluated, whereas state distributions provide the quantities pooled into the consensus.

4.2 Transition-Aware Credibility Mechanism

PosTR combines this structured context with market conditions, peer relations, and settled performance to estimate credibility relative to an equal consensus, then reweights state distributions and adaptively corrects the reference.

Space-conditioned credibility estimation. Credibility measures whether a trader's reported distribution is likely to outperform the equal consensus under the current context. After settlement, the realized return is mapped to an outcome distribution $\mathbf{y}_{t+1} \in \Delta^2$ over (Long, Flat, Short), with $u_{t+1} = r_{t+1}/(\sigma_t + |r_{t+1}|)$ and $\mathbf{y}_{t+1} = (\max(u_{t+1}, 0), 1 - |u_{t+1}|, \max(-u_{t+1}, 0))$. Using the multiclass Brier loss B (Brier, 1950; Gneiting & Raftery, 2007), we define the equal reference and the trader's relative-skill target as

$$\mathbf{P}_t^{\text{eq}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{P}_{i,t}, \tag{7}$$

$$z_{i,t+1} = B(\mathbf{P}_t^{\text{eq}}, \mathbf{y}_{t+1}) - B(\mathbf{P}_{i,t}, \mathbf{y}_{t+1}). \tag{8}$$

A positive $z_{i,t+1}$ indicates lower Brier loss than the equal consensus, rather than a guarantee of higher trading return. Before the next outcome is observed, credibility is estimated from the struc-

tured trader representation and its context:

$$c_{i,t} = \Pr\big(z_{i,t+1} > 0 \mid \boldsymbol{\phi}_{i,t}, \mathcal{H}_{t-1}\big)\,, \tag{9}$$

$$\boldsymbol{\phi}_{i,t} = \Phi(\zeta_{i,t}, \{\zeta_{j,t}\}_{j\neq i}, \mathbf{P}_t^{\mathrm{eq}}, \mathcal{I}_t, \mathcal{H}_{t-1})\,. \tag{10}$$

The 18-dimensional context combines transition location, signed movement, and state uncertainty with market conditions, peer relations, settled history, and trader roles. It includes movement magnitude, transition–opinion coherence, directional support, mean pairwise directional alignment, squared distance from the equal consensus, and their market and peer interactions. These features condition credibility on how a trader's position changes and relates to its peers, rather than on a global ranking of traders. A shared recursive Bayesian linear regression learner, initialized with zero mean and identity covariance, is trained only on previously settled $(\boldsymbol{\phi}_{i,\tau}, z_{i,\tau+1})_{\tau<t}$ pairs. The feature order and recursive updates are specified in the Supplementary Material.

Credibility-guided position consensus. Normalizing the credibility estimates gives

$$w_{i,t} = \frac{c_{i,t}}{\sum_{j=1}^{N} c_{j,t}}, \tag{11}$$

$$\mathbf{P}_t^{\mathrm{cred}} = \sum_{i=1}^{N} w_{i,t}\mathbf{P}_{i,t}. \tag{12}$$

Thus, transition and peer information influence the consensus through trader weights; the quantities being pooled remain the state distributions, not the transition coordinates.

Adaptive credibility correction. Because learned credibility may become less reliable as market conditions change, PosTR controls its influence by comparing two shadow branches that execute $\mathbf{P}_t^{\mathrm{eq}}$ and $\mathbf{P}_t^{\mathrm{cred}}$ under the same cost model. Let $\ell_\tau = \log(1 + g_\tau^{\mathrm{cred}}) - \log(1 + g_\tau^{\mathrm{eq}})$ be the settled after-cost log-growth advantage of the credibility branch on day τ. Its recent advantage determines the correction strength:

$$L_t = \sum_{\tau<t} \rho^{\,t-1-\tau}\ell_\tau, \qquad \lambda_t = \frac{1}{1+\exp(-\eta L_t)}, \tag{13}$$

$$\mathbf{P}_t^{*} = \mathbf{P}_t^{\mathrm{eq}} + \lambda_t\big(\mathbf{P}_t^{\mathrm{cred}} - \mathbf{P}_t^{\mathrm{eq}}\big)\,, \tag{14}$$

where $\rho \in (0, 1]$ discounts older settlements and $\eta > 0$ sets how sharply λ_t responds; $L_1 = 0$ gives $\lambda_1 = \frac{1}{2}$. Unlike a cumulative wealth share, which stays near $\frac{1}{2}$ whenever the two branches end with similar wealth, the discount and temperature let λ_t track which branch is currently more reliable. We use $\rho = 0.95$ and $\eta = 20$ throughout. Credibility controls relative trader influence; λ_t controls the strength of the correction to the equal reference.

4.3 PosTR Overall Framework

Figure 3 summarizes the sequential procedure. At step t, each trader's current output and retained target position form $\zeta_{i,t}$, and the final distribution determines the bounded system position:

$$x_t^* = P_t^*(L) - P_t^*(S) \in [-1, 1]. \tag{15}$$

The executable adjustment is $x_t^* - x_{t-1}^*$. Here, $s_{i,t-1} \to s_{i,t}$ describes a trader's target-position transition, whereas $x_{t-1}^* \to x_t^*$ describes the system's executed position change. Only after r_{t+1} is observed are the outcome distribution constructed and the credibility learner and the advantage L_t updated. Each branch uses its decoded position, realized return, and turnover cost to compute its net return; these updates affect subsequent decisions only.

5 Experiments

5.1 Experimental Setup

Data and Evaluation Protocol. We run daily single-asset backtests on eight U.S. equities from September 2, 2025, to August 31, 2026: AAPL (consumer electronics), NVDA (semiconductors),

Table 1: Overall trading performance on eight U.S. equities across diverse industries. ↑ indicates that higher values are better, while ↓ indicates that lower values are better.

Category	Method	AAPL				GOOGL				NVDA				TSLA			
		CR(%)↑	SR↑	MDD(%)↓	AV(%)↓	CR(%)↑	SR↑	MDD(%)↓	AV(%)↓	CR(%)↑	SR↑	MDD(%)↓	AV(%)↓	CR(%)↑	SR↑	MDD(%)↓	AV(%)↓
Market	B&H	38.02	**1.43**	13.80	25.08	61.02	1.63	21.05	32.71	29.45	0.87	20.21	37.99	11.72	0.47	39.10	46.76
Technical Indicator	MACD	23.37	1.28	10.43	17.71	32.51	1.29	17.17	24.05	-10.86	-0.35	26.72	24.58	31.06	1.05	22.71	30.29
	KDJ+RSI	13.49	0.94	11.32	**14.69**	22.05	**2.34**	**1.41**	**8.71**	11.17	0.65	14.19	**19.10**	14.67	1.09	**7.03**	**13.41**
	SMA	17.32	0.85	18.12	21.76	50.14	1.70	14.03	25.94	-2.89	0.06	20.57	31.72	-3.25	0.05	32.16	31.72
Traditional Quantitative	ZMR	19.62	1.25	**10.14**	15.29	15.79	0.79	15.67	21.44	45.47	**1.70**	**11.21**	23.78	-10.14	-0.16	28.49	33.12
	MV	15.10	0.80	12.15	20.23	52.30	1.77	17.20	25.73	2.42	0.23	25.94	30.68	6.98	0.37	28.14	30.93
	TSM	13.67	0.74	13.26	20.30	49.12	1.68	18.78	25.80	0.52	0.17	24.80	30.60	2.69	0.24	31.02	31.94
LLM-Based Trading	TradingAgents	15.78	0.88	13.73	18.72	12.59	0.69	15.29	20.19	-1.56	0.03	26.01	21.78	11.35	0.84	10.47	13.97
	FinMem	17.60	0.79	15.08	24.25	42.87	1.40	17.66	28.49	11.51	0.48	28.94	36.43	-0.39	0.21	40.01	42.64
Ours	PosTR	**38.37**	**1.43**	13.80	25.03	**82.55**	2.16	16.15	30.12	**68.61**	1.65	13.20	35.76	**50.44**	**1.23**	23.40	39.73

Category	Method	JPM				JNJ				CVX				COST			
		CR(%)↑	SR↑	MDD(%)↓	AV(%)↓	CR(%)↑	SR↑	MDD(%)↓	AV(%)↓	CR(%)↑	SR↑	MDD(%)↓	AV(%)↓	CR(%)↑	SR↑	MDD(%)↓	AV(%)↓
Market	B&H	21.05	0.97	15.47	22.32	52.66	2.36	10.96	18.85	32.38	1.32	20.82	23.57	1.13	0.16	16.57	19.95
Technical Indicator	MACD	1.27	0.16	10.49	16.95	12.12	0.92	13.79	13.47	-0.17	0.08	17.96	18.82	-30.00	-2.41	30.91	14.42
	KDJ+RSI	17.51	**1.46**	**4.72**	11.57	14.40	1.40	**5.18**	10.01	17.90	**2.62**	**1.09**	**6.38**	-0.17	0.04	13.20	11.54
	SMA	-1.22	0.02	18.15	17.68	53.14	2.62	7.57	16.86	13.85	0.89	14.89	16.18	-1.64	-0.05	13.54	14.10
Traditional Quantitative	ZMR	12.68	0.93	7.61	13.86	16.88	1.74	5.61	**9.22**	5.81	0.47	12.70	14.03	5.57	0.45	9.65	14.54
	MV	-2.47	-0.06	21.56	17.25	25.41	1.49	9.99	16.07	21.83	1.11	16.25	19.53	2.98	0.29	11.51	13.03
	TSM	-0.45	0.06	20.14	17.16	25.31	1.49	9.99	16.08	22.53	1.14	15.76	19.57	3.21	0.31	12.36	13.15
LLM-Based Trading	TradingAgents	-11.07	-1.10	16.35	**10.30**	4.22	0.50	6.04	9.23	0.16	0.07	7.25	11.39	-4.31	-0.66	**6.68**	**6.38**
	FinMem	4.99	0.33	18.96	21.50	43.74	2.11	11.70	18.01	25.29	1.13	20.47	22.22	-9.26	-0.44	18.83	18.47
Ours	PosTR	**34.22**	**1.46**	11.77	21.99	**60.77**	**2.67**	10.96	18.53	**49.72**	1.85	20.82	23.44	**11.20**	**0.64**	9.06	19.88

GOOGL (internet services), TSLA (electric vehicles), JPM (financial services), JNJ (healthcare), CVX (integrated energy), and COST (consumer retail). Three role-specialized traders produce fundamental, narrative, and momentum decisions with GPT-5.4-mini. PosTR and its matched ablations replay these frozen point-in-time outputs. B&H (buy-and-hold), technical indicators, quantitative rules, TradingAgents (Xiao et al., 2024), and FinMem (Yu et al., 2025) use the same window, next-day execution, and transaction cost, but not these frozen outputs.

A position earns the next trading day's return. Turnover is charged at 5 basis points per unit position change. Detailed experimental settings are given in the Supplementary Material.

Baselines and Evaluation Metrics. We report cumulative return (CR), Sharpe ratio (SR, annualized with a zero risk-free rate), maximum drawdown (MDD), and annualized volatility (AV), all on daily net returns; definitions are in the Supplementary Material. Standard deviations across equities measure cross-asset dispersion, not repeated-run uncertainty.

5.2 Overall Trading Performance

Table 1 is an end-to-end comparison under the shared window, execution timing, and cost. PosTR has the highest cumulative return on all eight equities and the highest Sharpe ratio on five. Cumulative return ranges from 11.20% to 82.55%, and the Sharpe ratio from 0.64 to 2.67. The cumulative-return gap versus the best non-PosTR baseline on each equity is positive for all eight equities and ranges from 0.35 to 23.14 percentage points. Maximum drawdown ranges from 9.06% to 23.40%. The table shows a higher cumulative-return point estimate on every equity, not uniformly lower risk. It does not identify the contribution of position-state aggregation, because the external systems do not use the frozen trader outputs. Figure 4 shows the same eight equities over time, against the three frozen specialists and B&H.

5.3 Component Analysis

Position-State Information. Table 2 fixes the credibility mechanism and varies only the input. Full state and Raw state share the same state pair and distribution, so their gap tests alternative feature constructions from the same state pair and distribution: $+4.8$ points of cumulative return, higher on six of eight equities. Relative to Endpoint, the transition path adds 6.4 points, the state distribution adds 10.3, and both together add 19.0, against 14.2 for Raw state. The joint gain exceeds the sum of the two single gains, so path and distribution are complementary rather than redundant. The

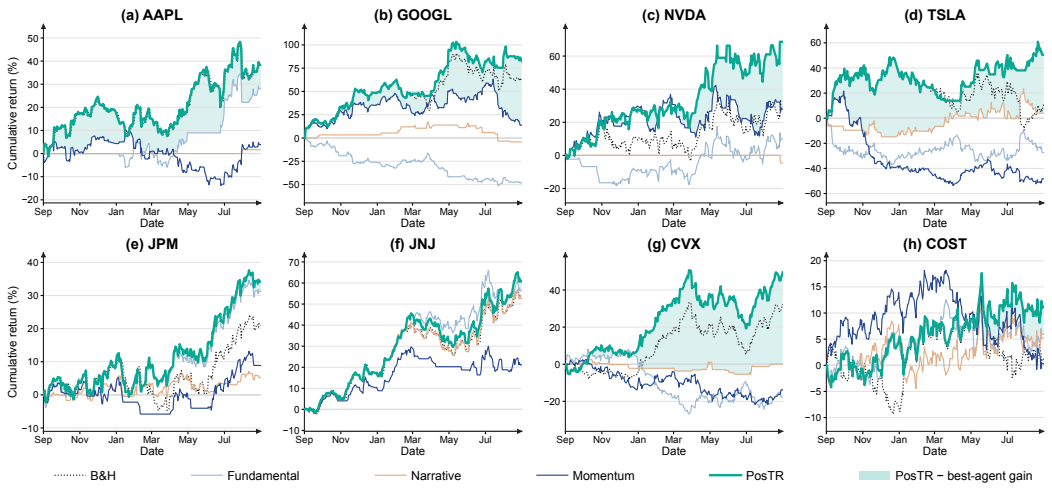

Figure 4: Cumulative returns of the fundamental, narrative, and momentum traders and of PosTR on the eight equities. The shaded band is PosTR's gap over the best specialist on each date; the dotted line is B&H.

Table 2: Position-state contrasts, credibility fixed. Δ: variant minus reference, mean over eight equities. n_{CR}: equities with higher CR.

Contrast	Reference	ΔCR	ΔSR	ΔMDD	ΔAV	n_{CR}
Same state pair and distribution						
Full state	Raw state	+4.8	+0.12	−1.1	−0.4	6/8
Relative to the current target						
Path	Endpoint	+6.4	+0.17	−1.3	−0.2	6/8
Semantics	Endpoint	+10.3	+0.29	−2.6	−2.9	7/8
Path + semantics	Endpoint	+19.0	+0.46	−3.7	−3.1	7/8
Raw state	Endpoint	+14.2	+0.34	−2.6	−2.7	7/8

Table 3: Credibility contrasts, input fixed at Full state. Δ and n_{CR} as in Table 2. Upper versus Equal; lower versus the row above.

Variant	ΔCR	ΔSR	ΔMDD	ΔAV	n_{CR}
Relative to Equal consensus					
History-only	+4.7	+0.11	−0.8	+0.9	5/8
Transition-aware	+13.1	+0.33	−2.1	+1.6	7/8
Full PosTR	+16.7	+0.44	−3.4	+1.0	7/8
Relative to the previous row					
+ Context	+8.4	+0.22	−1.3	+0.7	7/8
+ Correction	+3.6	+0.11	−1.3	−0.6	6/8

distribution also lowers volatility by 2.9 points, because it shrinks exposure when a trader's support is split.

Transition-Aware Credibility and Adaptive Correction. Table 3 fixes the input at Full state. History-only weights traders by settled performance, Transition-aware adds market, transition, and peer context, and Full PosTR adds the adaptive correction λ_t. Against Equal consensus, the three add 4.7, 13.1, and 16.7 points of cumulative return. The lower block isolates each step against the row above it: transition context adds 8.4 points beyond history, the largest single gain, and adaptive correction adds a further 3.6. Relative to Equal consensus, these gains come with higher volatility, but also with a higher Sharpe ratio and a lower maximum drawdown at every step; the correction gives back part of the volatility added by credibility.

5.4 Aggregation and Mechanism Analysis

Position Consensus across Specialized Traders. Table 4 asks whether PosTR's gain could come from selecting a specialist or counting votes instead of pooling. The best specialist changes across equities, and exactly two specialists agree on 30.8–72.4% of days. Leader finishes below the best specialist on every equity; Majority does so on seven of eight, averaging 9.25% and 4.83% against 20.91% for the best specialist in hindsight. PosTR exceeds all three specialists and both selection rules on every equity, by 2.99 (JPM) to 68.83 (GOOGL) points over the best specialist. Its excess over Leader accrues on split days (+16.89 points on average, positive on six of seven equities) and, to a lesser extent, on unanimous days (+7.51). Split days contribute more than unanimous days on every computed equity except TSLA, where the excess arises almost entirely on unanimous days.

Table 4: Consensus versus selection on the same frozen outputs. Majority holds the median target; Leader holds the best settled specialist. Excess is PosTR minus Leader, summed over split days (exactly two agree) or unanimous days; excess means are over the seven equities with computed decompositions. The last row counts equities where PosTR is higher.

	Specialists, CR (%)				Selection rules, CR (%)				Excess over Leader (pp)	
Equity	Fund.	Narr.	Mom.	Best	Majority	Leader	PosTR	Split days (%)	Split	Unanimous
AAPL	28.00	1.68	3.90	28.00	11.93	2.39	**38.37**	68.4	+21.18	−5.85
GOOGL	−48.16	−4.17	13.72	13.72	−15.17	4.20	**82.55**	48.0	+28.69	+12.25
NVDA	10.04	−4.67	27.95	27.95	3.72	19.76	**68.61**	66.8	+26.11	+8.34
TSLA	−27.37	3.97	−48.21	3.97	−16.42	−1.56	**50.44**	47.2	−0.50	+28.01
JPM	31.23	5.21	8.89	31.23	12.63	18.68	**34.22**	72.4	+15.66	−3.16
JNJ	56.38	52.58	21.15	56.38	55.04	46.30	**60.77**	41.6	+9.44	+0.12
CVX	−17.68	0.04	−13.77	0.04	−19.93	−11.83	**49.72**	45.2	+17.67	+12.87
COST	5.95	5.25	0.30	5.95	6.82	−3.93	**11.20**	30.8	–	–
Mean	4.80	7.49	1.74	20.91	4.83	9.25	**49.49**	52.6	+16.89	+7.51
PosTR > column	8/8	8/8	8/8	8/8	8/8	8/8				

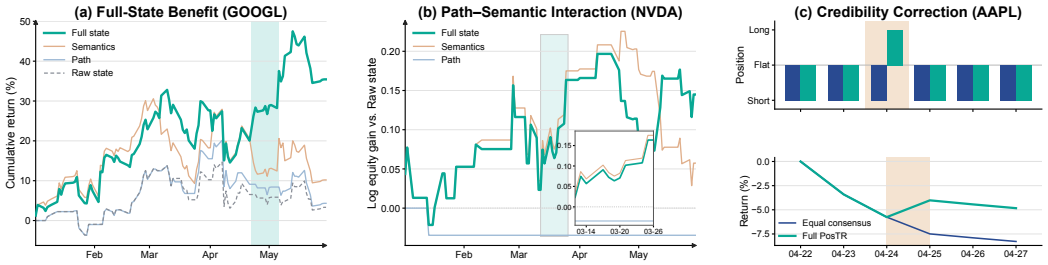

Figure 5: Mechanisms on a separate January–May 2026 replay. (a) GOOGL cumulative returns of the Table 2 inputs; shading marks where Full state separates. (b) NVDA log-equity gain over Raw state, March 12–26 enlarged. (c) AAPL Equal consensus and Full PosTR around April 24.

Representative Mechanisms. Figure 5 shows three cases in a separate January–May 2026 replay.

(1) **Full-state benefit.** On GOOGL, Full state ends at 35.43%, against 10.20% for Semantics, 4.33% for Path, and 3.32% for Raw state. The gap opens in late April and persists, so the benefit shapes the later trajectory rather than a single move.

(2) **Path–semantic interaction.** On NVDA, Path alone ends 0.035 below Raw state in log equity, while Full state ends 0.145 above it, more than Semantics alone (0.057). Path information that hurts in isolation helps once combined with semantics.

(3) **Credibility correction.** On AAPL on April 24, the only day the two positions differ, Equal consensus stays Short while Full PosTR goes Long. The next-day return of $+1.84\%$ yields a 3.68-point advantage, and the window ends at -4.83% against -8.28%.

6 Conclusion

We introduced PosTR, a position-transition reasoning framework that addresses *Decision Semantic Collapse* in multi-agent financial decisions. PosTR represents each trader in a shared position-state space, preserving transition paths and their implied market assessments while representing uncertainty over candidate positions. Space-conditioned credibility uses these representations and their peer relations to reweight state distributions into a system consensus, with settled performance controlling the correction strength. Experiments on eight equities across diverse industries demonstrate consistent improvements in profitability and risk-adjusted performance. These results suggest that reliable multi-agent financial decisions should reason over position-state transitions and space-conditioned aggregation rather than directly aggregate isolated trading actions.

AI USE STATEMENT

Large language models are the object of study in this work. The three role-specialized traders are generated by GPT-5.4-mini, the cross-backbone study in the Supplementary Material additionally uses DeepSeek-V3, Claude Sonnet 4.6, and Gemini 2.5 Flash, and the TradingAgents and FinMem baselines run on GPT-5.4-mini. All trader outputs were produced from point-in-time prompts and frozen before any replay. Beyond this role, we have not used generative AI tools to generate synthetic data, develop the conceptual framework, propose hypotheses, design the methodology or experiments, implement PosTR, or clean datasets; translation and qualitative data analysis are not applicable to this work. Additionally, we used generative AI tools to edit the manuscript for readability and length, to draft and shorten figure and table captions, to check consistency between the main text and the Supplementary Material, to discuss method definitions and mathematical claims, to review the interpretation of reported results against the experimental records, and to help write plotting and table-generation scripts. We have reviewed all AI-assisted work: edited text was checked against the method definitions and experimental records, and plotting and table scripts were checked against the replay outputs. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REFERENCES

Ralph Allan Bradley and Milton E Terry. Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952.

Glenn W. Brier. Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78(1):1–3, 1950. doi: 10.1175/1520-0493(1950)078⟨0001:VOFEIT⟩2.0.CO;2.

Chi-Min Chan, Weize Chen, Yusheng Su, Jianxuan Yu, Wei Xue, Shanghang Zhang, Jie Fu, and Zhiyuan Liu. Chateval: Towards better llm-based evaluators through multi-agent debate. In *International conference on learning representations*, volume 2024, pp. 9079–9093, 2024.

Justin Chen, Swarnadeep Saha, and Mohit Bansal. Reconcile: Round-table conference improves reasoning via consensus among diverse llms. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 7066–7085, 2024a.

Lili Chen, Kevin Lu, Aravind Rajeswaran, Kimin Lee, Aditya Grover, Misha Laskin, Pieter Abbeel, Aravind Srinivas, and Igor Mordatch. Decision transformer: Reinforcement learning via sequence modeling. *Advances in neural information processing systems*, 34:15084–15097, 2021.

Weize Chen, Yusheng Su, Jingwei Zuo, Cheng Yang, Chenfei Yuan, Chi-Min Chan, Heyang Yu, Yaxi Lu, Yi-Hsin Hung, Chen Qian, Yujia Qin, Xin Cong, Ruobing Xie, Zhiyuan Liu, Maosong Sun, and Jie Zhou. AgentVerse: Facilitating multi-agent collaboration and exploring emergent behaviors. In *Proceedings of the Twelfth International Conference on Learning Representations*, 2024b. URL https://openreview.net/forum?id=EHg5GDnyq1.

Hyeong Kyu Choi, Jerry Zhu, and Sharon Li. Debate or vote: Which yields better decisions in multi-agent large language models? *Advances in Neural Information Processing Systems*, 38: 101732–101764, 2026.

Abhishek Das, Théophile Gervet, Joshua Romoff, Dhruv Batra, Devi Parikh, Mike Rabbat, and Joelle Pineau. Tarmac: Targeted multi-agent communication. In *International Conference on machine learning*, pp. 1538–1546. PMLR, 2019.

Han Ding, Yinheng Li, Junhao Wang, Hang Chen, Doudou Guo, and Yunbai Zhang. Large language model agent in financial trading: A survey. In *Proceedings of the 2026 9th International Conference on Computers in Management and Business*, pp. 76–83, 2026.

Yifei Dong, Fengyi Wu, Kunlin Zhang, Yilong Dai, Sanjian Zhang, Wanghao Ye, Sihan Chen, and Zhi-Qi Cheng. Large language model agents in finance: A survey bridging research, practice, and real-world deployment. In *EMNLP (Findings)*, pp. 17889–17907, 2025.

Yilun Du, Shuang Li, Antonio Torralba, Joshua B Tenenbaum, and Igor Mordatch. Improving factuality and reasoning in language models through multiagent debate. *arXiv preprint arXiv:2305.14325*, 2023.

Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139, 1997.

Christian Genest and James V Zidek. Combining probability distributions: A critique and an annotated bibliography. *Statistical Science*, 1(1):114–135, 1986.

Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007. doi: 10.1198/016214506000001437.

Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q Weinberger. On calibration of modern neural networks. In *International conference on machine learning*, pp. 1321–1330. PMLR, 2017.

Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2):169–192, 2007.

Sirui Hong, Mingchen Zhuge, Jonathan Chen, Xiawu Zheng, Yuheng Cheng, Jinlin Wang, Ceyao Zhang, Zili Wang, Steven Ka Shing Yau, Zijuan Lin, Liyang Zhou, Chenyu Ran, Lingfeng Xiao, Chenglin Wu, and Jürgen Schmidhuber. MetaGPT: Meta programming for a multi-agent collaborative framework. In *Proceedings of the Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=VtmBAGCN7o.

Jen-tse Huang, Jiaxu Zhou, Tailin Jin, Xuhui Zhou, Zixi Chen, Wenxuan Wang, Youliang Yuan, Michael R Lyu, and Maarten Sap. On the resilience of llm-based multi-agent collaboration with faulty agents. *arXiv preprint arXiv:2408.00989*, 2024.

Michael Janner, Qiyang Li, and Sergey Levine. Offline reinforcement learning as one big sequence modeling problem. *Advances in neural information processing systems*, 34:1273–1286, 2021.

Dongfu Jiang, Xiang Ren, and Bill Yuchen Lin. Llm-blender: Ensembling large language models with pairwise ranking and generative fusion. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 14165–14178, 2023.

Meelis Kull, Miquel Perello Nieto, Markus Kängsepp, Telmo Silva Filho, Hao Song, and Peter Flach. Beyond temperature scaling: Obtaining well-calibrated multi-class probabilities with dirichlet calibration. *Advances in neural information processing systems*, 32, 2019.

Guohao Li, Hasan Hammoud, Hani Itani, Dmitrii Khizbullin, and Bernard Ghanem. Camel: Communicative agents for" mind" exploration of large language model society. *Advances in neural information processing systems*, 36:51991–52008, 2023.

Haohang Li, Yupeng Cao, Yangyang Yu, Shashidhar Reddy Javaji, Zhiyang Deng, Yueru He, Yuechen Jiang, Zining Zhu, K. P. Subbalakshmi, Jimin Huang, Lingfei Qian, Xueqing Peng, Jordan W. Suchow, and Qianqian Xie. INVESTORBENCH: A benchmark for financial decision-making tasks with LLM-based agent. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics*, pp. 2509–2525, 2025a. doi: 10.18653/v1/2025.acl-long.126.

Xiangyu Li, Yawen Zeng, Xiaofen Xing, Jin Xu, and Xiangmin Xu. Hedgeagents: A balanced-aware multi-agent financial trading system. In *Companion Proceedings of the ACM on Web Conference 2025*, pp. 296–305, 2025b.

Xiangyu Li, Yawen Zeng, Xiaofen Xing, Jin Xu, and Xiangmin Xu. Quantagents: Towards multi-agent financial system via simulated trading. *arXiv preprint arXiv:2510.04643*, 2025c.

Yuan Li, Bingqiao Luo, Qian Wang, Nuo Chen, Xu Liu, and Bingsheng He. Cryptotrade: A reflective llm-based agent to guide zero-shot cryptocurrency trading. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 1094–1106, 2024.

Nick Littlestone and Manfred K Warmuth. The weighted majority algorithm. *Information and computation*, 108(2):212–261, 1994.

Bijia Liu and Ronghao Dang. Finpos: A position-aware trading agent system for real financial markets. *arXiv preprint arXiv:2510.27251*, 2025.

Haipeng Luo and Robert E Schapire. Achieving all with no parameters: Adanormalhedge. In *Conference on Learning Theory*, pp. 1286–1304. PMLR, 2015.

Kunihiro Miyazaki, Takanobu Kawahara, Stephen Roberts, and Stefan Zohren. Toward expert investment teams: A multi-agent llm system with fine-grained trading tasks. *arXiv preprint arXiv:2602.23330*, 2026.

Chunjiang Mu, Ya Zeng, Qiaosheng Zhang, Kun Shao, Chen Chu, Hao Guo, Danyang Jia, Zhen Wang, and Shuyue Hu. Adaptive theory of mind for llm-based multi-agent coordination. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 29608–29616, 2026.

Dimitris N. Politis and Joseph P. Romano. The stationary bootstrap. *Journal of the American Statistical Association*, 89(428):1303–1313, 1994. doi: 10.1080/01621459.1994.10476870.

Siyue Ren, Wanli Fu, Xinkun Zou, Chen Shen, Yi Cai, Chen Chu, Zhen Wang, and Shuyue Hu. Reputation as a solution to cooperation collapse in llm-based mass. *arXiv preprint arXiv:2505.05029*, 2025.

Murat Sensoy, Lance Kaplan, and Melih Kandemir. Evidential deep learning to quantify classification uncertainty. *Advances in neural information processing systems*, 31, 2018.

Noah Shinn, Federico Cassano, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. Reflexion: Language agents with verbal reinforcement learning. *Advances in neural information processing systems*, 36:8634–8652, 2023.

Shuo Sun, Molei Qin, Wentao Zhang, Haochong Xia, Chuqiao Zong, Jie Ying, Yonggang Xie, Lingxuan Zhao, Xinrun Wang, and Bo An. Trademaster: A holistic quantitative trading platform empowered by reinforcement learning. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

Feng Tian, Flora D Salim, and Hao Xue. Tradinggroup: A multi-agent trading system with self-reflection and data-synthesis. *arXiv preprint arXiv:2508.17565*, 2025.

Sharlin Utke, Jeremie Houssineau, and Giovanni Montana. Investigating relational state abstraction in collaborative marl. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 20947–20955, 2025.

Junlin Wang, Jue Wang, Ben Athiwaratkun, Ce Zhang, and James Y Zou. Mixture-of-agents enhances large language model capabilities. In *International Conference on Learning Representations*, volume 2025, pp. 33944–33963, 2025.

Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou. Self-consistency improves chain of thought reasoning in language models. *arXiv preprint arXiv:2203.11171*, 2022.

Zhicheng Wang, Biwei Huang, Shikui Tu, Kun Zhang, and Lei Xu. Deeptrader: A deep reinforcement learning approach for risk-return balanced portfolio management with market conditions embedding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 643–650, 2021. doi: 10.1609/aaai.v35i1.16144.

Muning Wen, Jakub Kuba, Runji Lin, Weinan Zhang, Ying Wen, Jun Wang, and Yaodong Yang. Multi-agent reinforcement learning is a sequence modeling problem. *Advances in Neural Information Processing Systems*, 35:16509–16521, 2022.

Qingyun Wu, Gagan Bansal, Jieyu Zhang, Yuxuan Wu, Beibin Li, Erkang Zhu, Li Jiang, Xiaoyun Zhang, Shaokun Zhang, Jiale Liu, Ahmed Hassan Awadallah, Ryen W. White, Doug Burger, and Chi Wang. AutoGen: Enabling next-gen LLM applications via multi-agent conversation. *arXiv preprint arXiv:2308.08155*, 2023. URL `https://arxiv.org/abs/2308.08155`.

Yihan Xia, Taotao Wang, Shengli Zhang, Zhangyuhua Weng, Bin Cao, and Soung Chang Liew. Hivemind: Contribution-guided online prompt optimization of llm multi-agent systems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 29767–29774, 2026.

Yijia Xiao, Edward Sun, Di Luo, and Wei Wang. Tradingagents: Multi-agents llm financial trading framework. *arXiv preprint arXiv:2412.20138*, 2024.

Guojun Xiong, Zhiyang Deng, Keyi Wang, Yupeng Cao, Haohang Li, Yangyang Yu, Xueqing Peng, Mingquan Lin, Kaleb E Smith, Xiao-Yang Liu, et al. Flag-trader: Fusion llm-agent with gradient-based reinforcement learning for financial trading. In *Findings of the Association for Computational Linguistics: ACL 2025*, pp. 13921–13934, 2025.

Xie Yi, Zhanke Zhou, Chentao Cao, Qiyu Niu, Tongliang Liu, and Bo Han. From debate to equilibrium: Belief-driven multi-agent llm reasoning via bayesian nash equilibrium. *arXiv preprint arXiv:2506.08292*, 2025.

Yangyang Yu, Zhiyuan Yao, Haohang Li, Zhiyang Deng, Yuechen Jiang, Yupeng Cao, Zhi Chen, Jordan W Suchow, Zhenyu Cui, Rong Liu, et al. Fincon: A synthesized llm multi-agent system with conceptual verbal reinforcement for enhanced financial decision making. *Advances in Neural Information Processing Systems*, 37:137010–137045, 2024.

Yangyang Yu, Haohang Li, Zhi Chen, Yuechen Jiang, Yang Li, Jordan W Suchow, Denghui Zhang, and Khaldoun Khashanah. Finmem: A performance-enhanced llm trading agent with layered memory and character design. *IEEE Transactions on Big Data*, 2025.

Shaokun Zhang, Ming Yin, Jieyu Zhang, Jiale Liu, Zhiguang Han, Jingyang Zhang, Beibin Li, Chi Wang, Huazheng Wang, Yiran Chen, et al. Which agent causes task failures and when? on automated failure attribution of llm multi-agent systems. *arXiv preprint arXiv:2505.00212*, 2025.

Wentao Zhang, Lingxuan Zhao, Haochong Xia, Shuo Sun, Jiaze Sun, Molei Qin, Xinyi Li, Yuqing Zhao, Yilei Zhao, Xinyu Cai, et al. A multimodal foundation agent for financial trading: Tool-augmented, diversified, and generalist. In *Proceedings of the 30th acm sigkdd conference on knowledge discovery and data mining*, pp. 4314–4325, 2024.

Supplementary Material

PosTR: Position Transition Reasoning for Multi-Agent Financial Decisions

OVERVIEW

This supplement supports the main paper along five lines.

- **Formal foundations** (Sections A–B). We fix notation, formalize *Decision Semantic Collapse*, and prove that the Position-State Space is lossless, that PosTR always outputs a valid bounded position, and that every decision is causal.
- **Complete specification** (Section C). We give every quantity omitted from the main text: the nine transitions, the outcome distribution, the exact 18-dimensional credibility context, the recursive Bayesian learner, the adaptive correction, and the full online procedure.
- **Empirical evidence for the problem** (Section D). Using the frozen trader outputs, we show that Decision Semantic Collapse occurs in practice: identical action labels hide several distinct transition paths, and identical transitions hide different state distributions.
- **Experimental protocol** (Sections E–G). We document the data, trader roles, prompts, output validation, baselines, metrics, and the exact information boundary of every matched ablation.
- **Extended analyses and reproducibility** (Sections H–J). We report per-asset ablations, robustness checks, mechanism diagnostics, audit invariants, and limitations.

All experiments follow the main-paper protocol: eight U.S. equities across diverse industries, daily decisions from September 2, 2025, to August 31, 2026, three role-specialized traders whose GPT-5.4-mini outputs are frozen before replay, next-trading-day execution, and 5 basis points per unit of turnover.

A NOTATION

Table S1: Notation used throughout the main paper and this supplement.

Symbol	Meaning
N	Number of role-specialized traders ($N = 3$: fundamental, narrative, momentum)
t	Decision step (trading day)
$\mathcal{S} = \{L, F, S\}$	Target-position states, ordered $S \prec F \prec L$
$\psi(\cdot)$	Ordered encoding $\psi(S) = -1,\ \psi(F) = 0,\ \psi(L) = 1$
$s_{i,t}$	Encoded target position of trader i at step t
$\tau_{i,t} = (s_{i,t-1}, s_{i,t})$	Position transition of trader i
$m_{i,t},\ v_{i,t}$	Transition location and signed target-position change (Eq. S3)
$\mathbf{P}_{i,t} = (p^L_{i,t}, p^F_{i,t}, p^S_{i,t})$	State distribution over candidate target positions, $\mathbf{P}_{i,t} \in \Delta^2$
$e_{i,t}$	Supporting evidence text (validated, not used in aggregation)
$\zeta_{i,t} = (m_{i,t}, v_{i,t}, \mathbf{P}_{i,t})$	Structured trader representation
$\mathbf{P}^{\mathrm{eq}}_t,\ \mathbf{P}^{\mathrm{cred}}_t,\ \mathbf{P}^*_t$	Equal, credibility-weighted, and final consensus distributions
$\boldsymbol{\phi}_{i,t} \in \mathbb{R}^{18}$	Space-conditioned credibility context
$c_{i,t},\ w_{i,t}$	Credibility (posterior probability of positive relative skill) and normalized weight
$z_{i,t+1}$	Settled relative skill against the equal consensus
$\mathbf{y}_{t+1}$	Outcome distribution built from the realized return
λ_t	Adaptive correction strength
$\ell_\tau,\ L_t$	Daily log-growth advantage of the credibility branch and its discounted sum
$\rho,\ \eta$	Discount and temperature of the adaptive correction ($\rho = 0.95$, $\eta = 20$)
$x^*_t \in [-1, 1]$	Executed system position
r_{t+1}	Next-trading-day asset return
σ_t	Ex-ante daily volatility (10-day realized, known at t)
γ	Transaction cost per unit turnover ($\gamma = 0.0005$)
$\mathcal{I}_t,\ \mathcal{H}_{t-1}$	Market information at t; outcomes settled before t

A trader's *target position* $s_{i,t}$ is the exposure it prefers; the *system position* x_t^* is the single exposure produced after aggregation. The trader transition $s_{i,t-1} \to s_{i,t}$ and the system adjustment $x_{t-1}^* \to x_t^*$ are therefore different objects: the former is an input to aggregation, the latter its executable consequence.

B Theoretical Properties

This section formalizes the claims made informally in Sections 3 and 4 of the main paper.

B.1 Decision Semantic Collapse

Definition 1 (Directional action map)**.** Under the directional action convention, an action label is the sign of the target-position change,

$$\alpha(\tau) = \begin{cases} \textsc{Buy}, & s_t > s_{t-1}, \\ \textsc{Hold}, & s_t = s_{t-1}, \\ \textsc{Sell}, & s_t < s_{t-1}. \end{cases} \tag{S1}$$

Proposition 1 (Action labels are many-to-one)**.** *α maps the nine ordered transitions on $\mathcal{S}$ onto three labels with fibres*

$$\alpha^{-1}(\textsc{Buy}) = \{SF,\ FL,\ SL\}, \qquad \alpha^{-1}(\textsc{Hold}) = \{SS,\ FF,\ LL\},$$
$$\alpha^{-1}(\textsc{Sell}) = \{FS,\ LF,\ LS\}.$$

Each fibre contains three transitions that differ in origin, destination, or span.

Proof. The sign of $s_t - s_{t-1}$ partitions $\mathcal{S} \times \mathcal{S}$ into three classes. Enumerating the nine pairs gives the stated fibres, and within each fibre no two pairs share both endpoints. □

Corollary 1 (Action-only aggregation cannot recover transitions)**.** *Let G be any aggregation rule whose trader inputs are only the action labels $\{\alpha(\tau_{i,t})\}_{i=1}^N$ (together with any quantity that does not depend on the traders' target positions). For any two trader profiles $\{\tau_{i,t}\}$ and $\{\tau'_{i,t}\}$ with $\alpha(\tau_{i,t}) = \alpha(\tau'_{i,t})$ for all i, G returns the same output, even when the profiles differ in every trader's origin, destination, or span.*

Proof. G factors through α, and by Proposition 1 distinct transition profiles can have identical label profiles. □

Corollary 1 is the formal content of the statement that information lost before aggregation cannot be recovered from the action label alone. For instance, a profile in which all three traders move $S \to F$ (a pessimistic view softening to neutral) and a profile in which all three move $F \to L$ (a neutral view turning optimistic) are indistinguishable to G, yet they call for different system exposures.

B.2 Losslessness of the Position-State Space

Proposition 2 (Invertibility)**.** *The map $(s_{t-1}, s_t) \mapsto (m, v) = \left(\frac{s_{t-1}+s_t}{2}, \frac{s_t - s_{t-1}}{2}\right)$ is a bijection from $\{-1, 0, 1\}^2$ onto the nine lattice points listed in Table S2, with inverse $s_{t-1} = m - v$ and $s_t = m + v$. Moreover, $\operatorname{sign}(v)$ recovers the action label, $|v|$ recovers the span of the transition, and $|m| + |v| \le 1$.*

Proof. The map is linear with matrix $\frac{1}{2}\left(\begin{smallmatrix} 1 & 1 \\ -1 & 1 \end{smallmatrix}\right)$, whose determinant is $\frac{1}{2} \neq 0$, so it is injective; its image on the nine integer pairs is the set in Table S2. The inverse follows by adding and subtracting the two coordinates. Since $s_t - s_{t-1} = 2v$, $\operatorname{sign}(v) = \operatorname{sign}(s_t - s_{t-1})$ equals the action label in Eq. (S1). Finally, $|m| + |v| = \max(|s_{t-1}|, |s_t|) \le 1$. □

Proposition 2 makes precise that the Position-State Space retains everything the action label retains (through $\operatorname{sign}(v)$) and, in addition, the origin, destination, and span that the label discards. The factorization into (m, v) adds no information relative to the raw pair; its value is an inductive bias that places transitions sharing a direction on a common axis and transitions sharing a location on another, which is what the Full-state versus Raw-state ablation in the main paper tests.

B.3 Validity and Boundedness of the Output

Proposition 3 (Valid consensus and bounded position). *For every t: (i) $w_{i,t} > 0$ and $\sum_i w_{i,t} = 1$; (ii) $\lambda_t \in (0,1)$; (iii) $\mathbf{P}_t^{\mathrm{eq}}, \mathbf{P}_t^{\mathrm{cred}}, \mathbf{P}_t^* \in \Delta^2$; (iv) $x_t^* \in [-1,1]$; and (v) the per-step turnover satisfies $|x_t^* - x_{t-1}^*| \le 2$.*

Proof. (i) $c_{i,t} = \Phi(\cdot) \in (0,1)$ because the predictive variance in Eq. S12 is strictly greater than one, so normalization is well defined. (ii) Branch net returns satisfy $1 + g > 0$ whenever $|x| \le 1$, $|r_{t+1}| < 1 - 2\gamma$, and $\gamma < \frac{1}{2}$, so every ℓ_τ and hence L_t is finite, and $\lambda_t = 1/(1 + e^{-\eta L_t}) \in (0,1)$. (iii) $\mathbf{P}_t^{\mathrm{eq}}$ and $\mathbf{P}_t^{\mathrm{cred}}$ are convex combinations of points of Δ^2, and $\mathbf{P}_t^* = (1-\lambda_t)\mathbf{P}_t^{\mathrm{eq}} + \lambda_t \mathbf{P}_t^{\mathrm{cred}}$ is a convex combination of those two. (iv) For $\mathbf{P} \in \Delta^2$, $P(L) - P(S) \in [-1,1]$. (v) follows from (iv). □

Remark 1. Because $\mathbf{P}_t^*$ lies on the segment between $\mathbf{P}_t^{\mathrm{eq}}$ and $\mathbf{P}_t^{\mathrm{cred}}$, PosTR never moves farther from the equal reference than the credibility-weighted consensus does. The adaptive correction can only shrink, never amplify, the credibility-based revision, which limits the damage when learned credibility degrades.

B.4 Causality

Proposition 4 (No look-ahead). *Let $\mathcal{F}_t$ be the information set containing $\mathcal{I}_t$, the frozen trader outputs dated no later than t, and all returns r_τ with $\tau \le t$. Then x_t^* is $\mathcal{F}_t$-measurable.*

Proof. $\zeta_{i,t}$ uses $s_{i,t-1}$, $s_{i,t}$, and $\mathbf{P}_{i,t}$, all in $\mathcal{F}_t$. The context $\boldsymbol{\phi}_{i,t}$ uses market features dated t and peer representations at t. The learner state $(\boldsymbol{\beta}_{t-1}, \mathbf{C}_{t-1})$ is a deterministic function of settled pairs $(\boldsymbol{\phi}_{i,\tau}, z_{i,\tau+1})$ with $\tau + 1 \le t$, and $z_{i,\tau+1}$ depends on $r_{\tau+1}$ only. The advantage L_t uses branch returns settled with r_τ, $\tau \le t$. Hence $c_{i,t}$, $w_{i,t}$, λ_t, $\mathbf{P}_t^*$, and x_t^* are all $\mathcal{F}_t$-measurable. The return r_{t+1} enters only through $z_{i,t+1}$ and W_t, which affect decisions at $t+1$ and later. □

B.5 Properties of the Relative-Skill Target

Proposition 5 (Bounded relative skill). *For any realized return, $\mathbf{y}_{t+1} \in \Delta^2$ and $z_{i,t+1} \in [-2,2]$.*

Proof. With $u = r/(\sigma + |r|) \in (-1,1)$, the three components of $\mathbf{y}$ are nonnegative and sum to $\max(u,0) + 1 - |u| + \max(-u,0) = 1$. The Brier loss between two points of Δ^2 lies in $[0,2]$, so the difference of two such losses lies in $[-2,2]$. □

The equal reference $\mathbf{P}^{\mathrm{eq}}$ itself contains the evaluated trader, so changing $\mathbf{P}_{i,t}$ also changes the reference. Strict propriety of the Brier score therefore does not imply that the conditional expectation of $\mathbf{y}_{t+1}$ uniquely maximizes $\mathbb{E}[z_{i,t+1}]$. Proposition 5 is used only as a bound: the target is comparable across assets with different volatility, and it is not a guarantee of higher trading return.

B.6 Computational Cost

Per decision step, forming $\zeta_{i,t}$ and the context is $O(N^2)$ because of the pairwise alignment term, computing $c_{i,t}$ is $O(Nd^2)$ with $d = 18$, and the recursive update is $O(Nd^2)$. With $N = 3$ the aggregation cost is negligible relative to trader generation; all LLM calls occur before freezing, and replay requires no further model calls.

C Complete PosTR Specification

C.1 Trader Output Contract

For trader i and date t, the frozen output is

$$\mathcal{O}_{i,t} = (s_{i,t}, \mathbf{P}_{i,t}, e_{i,t}), \qquad \mathbf{P}_{i,t} \in \Delta^2, \tag{S2}$$

with $\mathbf{P}_{i,t}$ ordered as (Long, Flat, Short). A record is accepted only if the probabilities are finite, nonnegative, and sum to one; $s_{i,t}$ is a mode of $\mathbf{P}_{i,t}$; and $e_{i,t}$ is nonempty. Invalid records are rejected, never repaired or imputed.

C.2 Position-State Transition

With the encoding $\psi(S) = -1$, $\psi(F) = 0$, $\psi(L) = 1$,

$$m_{i,t} = \frac{s_{i,t-1} + s_{i,t}}{2}, \qquad v_{i,t} = \frac{s_{i,t} - s_{i,t-1}}{2}. \tag{S3}$$

Table S2 enumerates the complete space together with the action label that each transition collapses to.

Table S2: The nine position-state transitions, their coordinates, and the action label each collapses to. Every label covers three transitions.

Origin	Target	m	v	Action label	Implied assessment
S	S	-1	0	Hold	pessimistic, unchanged
S	F	-0.5	0.5	Buy	pessimistic → neutral
S	L	0	1	Buy	pessimistic → optimistic (full reversal)
F	S	-0.5	-0.5	Sell	neutral → pessimistic
F	F	0	0	Hold	neutral, unchanged
F	L	0.5	0.5	Buy	neutral → optimistic
L	S	0	-1	Sell	optimistic → pessimistic (full reversal)
L	F	0.5	-0.5	Sell	optimistic → neutral
L	L	1	0	Hold	optimistic, unchanged

C.3 Distributional Quantities

From $\mathbf{P}_{i,t}$ PosTR derives the directional support $d_{i,t} = p^L_{i,t} - p^S_{i,t} \in [-1, 1]$ and the neutral support $p^F_{i,t}$. Two traders on the same transition can differ in both: one may move $S \to F$ with residual pessimism ($d < 0$), another with residual optimism ($d > 0$). The transition fixes the path; the distribution refines how strongly, and in which residual direction, the trader holds its current target.

C.4 Outcome Distribution and Relative Skill

After r_{t+1} is observed,

$$u_{t+1} = \frac{r_{t+1}}{\sigma_t + |r_{t+1}|}, \qquad \mathbf{y}_{t+1} = \big(\max(u_{t+1}, 0),\ 1 - |u_{t+1}|,\ \max(-u_{t+1}, 0)\big), \tag{S4}$$

where σ_t is the 10-day realized daily volatility known at t. A move that is small relative to σ_t places most mass on Flat; a large up- or down-move places most mass on Long or Short. With $B(\mathbf{p}, \mathbf{y}) = \sum_k (p_k - y_k)^2$,

$$\mathbf{P}^{\text{eq}}_t = \frac{1}{N} \sum_{i=1}^{N} \mathbf{P}_{i,t}, \qquad z_{i,t+1} = B(\mathbf{P}^{\text{eq}}_t, \mathbf{y}_{t+1}) - B(\mathbf{P}_{i,t}, \mathbf{y}_{t+1}). \tag{S5}$$

C.5 Space-Conditioned Credibility Context

The context $\boldsymbol{\phi}_{i,t} \in \mathbb{R}^{18}$ is listed in Table S3. Market terms are shared by all traders at t; representation and peer terms are trader-specific. All market quantities are bounded to $(-1, 1)$ by volatility normalization, so a single learner can be shared across traders and remains comparable across assets.

Let $\rho_t^{(k)}$ be the k-day return, $\delta_t^{(20)}$ the 20-day drawdown from the running peak, and σ_t the 10-day daily volatility. Define

$$h_t = \frac{\rho_t^{(20)}}{\sigma_t\sqrt{20} + |\rho_t^{(20)}|}, \qquad h_t^s = \frac{\rho_t^{(5)}}{\sigma_t\sqrt{5} + |\rho_t^{(5)}|}, \qquad \Delta h_t = h_t^s - h_t, \tag{S6}$$

$$\tilde{\sigma}_t = \frac{\sqrt{252}\,\sigma_t}{\sqrt{1 + 252\,\sigma_t^2}}, \qquad R_t = \frac{\delta_t^{(20)}}{\sigma_t\sqrt{20} + |\delta_t^{(20)}|}, \qquad D_t = \sqrt{\frac{1}{N}\sum_i (s_{i,t} - \bar{s}_t)^2}, \tag{S7}$$

and, for trader i,

$$a_{i,t} = \frac{1}{N-1}\sum_{j \neq i} d_{i,t} d_{j,t}, \qquad g_{i,t} = \left\|\mathbf{P}_{i,t} - \mathbf{P}_t^{\mathrm{eq}}\right\|_2^2, \qquad \kappa_{i,t} = d_{i,t}\, v_{i,t}. \tag{S8}$$

The coherence term $\kappa_{i,t}$ is positive when the trader's residual directional support agrees with the direction of its transition (e.g., a $F \to L$ move backed by $p^L > p^S$) and negative when the transition and the distribution pull in opposite directions.

Table S3: The 18-dimensional space-conditioned credibility context $\boldsymbol{\phi}_{i,t}$. The group column defines the masks used by the matched ablations in Section G.

No.	Feature	Meaning	Group
1	1	intercept	–
2	h_t	volatility-normalized 20-day trend	market
3	h_t^s	volatility-normalized 5-day trend	market
4	Δh_t	trend change (short minus long)	market
5	$\tilde{\sigma}_t$	bounded annualized volatility	market
6	D_t	cross-trader disagreement of target positions	peer
7	R_t	volatility-normalized drawdown	market
8	$d_{i,t}$	directional support $p^L - p^S$	distribution
9	$a_{i,t}$	mean pairwise directional alignment with peers	peer
10	$g_{i,t}$	squared distance from the equal consensus	distribution/peer
11	$\lvert v_{i,t}\rvert$	transition movement magnitude	transition
12	$\kappa_{i,t}$	transition–distribution coherence $d_{i,t}v_{i,t}$	transition × distribution
13	$\Delta h_t\, d_{i,t}$	trend-change interaction	market × distribution
14	$D_t\, d_{i,t}$	disagreement interaction	peer × distribution
15	$R_t\, d_{i,t}$	drawdown–direction interaction	market × distribution
16	$R_t\, a_{i,t}$	drawdown–peer interaction	market × peer
17	$I_i^{(1)}$	fundamental-role indicator	role
18	$I_i^{(2)}$	narrative-role indicator (momentum is reference)	role

The context deliberately contains no trader identity beyond two role indicators and no asset identity. Credibility is therefore a function of *how* a trader's position changes and relates to its peers under the current market, not a fixed ranking of traders.

C.6 Recursive Bayesian Credibility Learner

A single linear-Gaussian model $z = \boldsymbol{\phi}^\top \boldsymbol{\beta} + \varepsilon$, $\varepsilon \sim \mathcal{N}(0,1)$, is shared by all traders, with prior $\boldsymbol{\beta}_0 = \mathbf{0}$, $\mathbf{C}_0 = \mathbf{I}$. Each settled pair $(\boldsymbol{\phi}, z)$ updates

$$\mathbf{k} = \frac{\mathbf{C}\boldsymbol{\phi}}{1 + \boldsymbol{\phi}^\top \mathbf{C}\boldsymbol{\phi}}, \tag{S9}$$

$$\boldsymbol{\beta} \leftarrow \boldsymbol{\beta} + \mathbf{k}\,(z - \boldsymbol{\phi}^\top \boldsymbol{\beta}), \tag{S10}$$

$$\mathbf{C} \leftarrow \mathbf{C} - \mathbf{k}\,\boldsymbol{\phi}^\top \mathbf{C}. \tag{S11}$$

The N pairs settled on a date are applied sequentially in fixed role order. Before settlement, credibility is the one-step predictive probability that the next relative-skill residual is positive, as defined

in the main paper. Under the linear-Gaussian model $z = \boldsymbol{\phi}^\top \boldsymbol{\beta} + \varepsilon$ with $\varepsilon \sim \mathcal{N}(0,1)$, this is

$$c_{i,t} = \Pr\big(z_{i,t+1} > 0 \mid \boldsymbol{\phi}_{i,t}, \mathcal{H}_{t-1}\big) = \Phi\left(\frac{\boldsymbol{\phi}_{i,t}^\top \boldsymbol{\beta}_{t-1}}{\sqrt{1 + \boldsymbol{\phi}_{i,t}^\top \mathbf{C}_{t-1} \boldsymbol{\phi}_{i,t}}}\right), \qquad w_{i,t} = \frac{c_{i,t}}{\sum_j c_{j,t}}, \tag{S12}$$

where Φ is the standard normal CDF. The observation-noise term in the denominator distinguishes this quantity from the posterior probability that the conditional mean alone is positive. At initialization $c_{i,1} = \frac{1}{2}$ for every trader, so PosTR starts exactly at the equal consensus. Because $c_{i,t}$ divides the posterior mean by the predictive standard deviation, a context seen rarely (large posterior variance) yields credibility near $\frac{1}{2}$ even if its point estimate is extreme; this is the mechanism that keeps early and unusual decisions close to the equal reference.

C.7 Credibility-Guided Consensus and Adaptive Correction

$$\mathbf{P}_t^{\mathrm{cred}} = \sum_{i=1}^{N} w_{i,t} \mathbf{P}_{i,t}, \qquad \mathbf{P}_t^{*} = \mathbf{P}_t^{\mathrm{eq}} + \lambda_t \big(\mathbf{P}_t^{\mathrm{cred}} - \mathbf{P}_t^{\mathrm{eq}}\big). \tag{S13}$$

The two shadow branches execute the decoded positions of $\mathbf{P}_t^{\mathrm{eq}}$ and $\mathbf{P}_t^{\mathrm{cred}}$ under the same cost model. With g_τ^{eq} and g_τ^{cred} their net returns settled on day τ,

$$\ell_\tau = \log \frac{1 + g_\tau^{\mathrm{cred}}}{1 + g_\tau^{\mathrm{eq}}}, \qquad L_t = \rho\, L_{t-1} + \ell_{t-1}, \qquad \lambda_t = \frac{1}{1 + \exp(-\eta L_t)}, \tag{S14}$$

with $L_1 = 0$, so $\lambda_1 = \frac{1}{2}$. Unrolling the recursion gives $L_t = \sum_{\tau<t} \rho^{\,t-1-\tau} \ell_\tau$, the discounted log-wealth ratio of the two branches.

Why a discount and a temperature. With $\eta = \rho = 1$, $\eta L_t = \log(W_{t-1}^{\mathrm{cred}} / W_{t-1}^{\mathrm{eq}})$ and Eq. (S14) reduces to the cumulative wealth share $W_{t-1}^{\mathrm{cred}} / (W_{t-1}^{\mathrm{eq}} + W_{t-1}^{\mathrm{cred}})$. That share moves very little: two branches whose cumulative returns differ by ten points over a year give $\lambda \approx 0.52$, so the final consensus stays near the midpoint of the two branches and cannot exceed the better one. A daily advantage ℓ_τ is typically a few tenths of a percent; with $\rho = 0.95$ (effective memory of about 20 trading days) L_t is of order 10^{-2}, and $\eta = 20$ maps it to λ_t roughly in $[0.35, 0.75]$. λ_t therefore leans toward whichever branch has been more reliable over the recent month while never fully abandoning either, and the final consensus can outperform both branches when their relative reliability persists over weeks. Section H.10 reports sensitivity to both parameters, including the undiscounted special case.

The system position and its settlement are

$$x_t^* = P_t^*(L) - P_t^*(S), \qquad q_t = |x_t^* - x_{t-1}^*|, \qquad g_{t+1} = x_t^*\, r_{t+1} - \gamma\, q_t, \tag{S15}$$

with $x_0^* = 0$ and $\gamma = 0.0005$. Each branch computes its net return with the same formula applied to its own decoded position. The executed exposure x_t^* is continuous in $[-1, 1]$ and need not coincide with any specialist's categorical target: even on a unanimous day the consensus can differ in magnitude from ± 1, which is the source of the unanimous-day excess in main-paper Table 4.

C.8 Complete Online Procedure

Algorithm S1 PosTR online decision and settlement

Require: Frozen outputs $\{\mathcal{O}_{i,t}\}$; point-in-time market features
$x_0^* \leftarrow 0$; $\boldsymbol{\beta} \leftarrow \mathbf{0}$; $\mathbf{C} \leftarrow \mathbf{I}$; $L \leftarrow 0$
for decision dates $t = 1, \ldots, T$ in chronological order **do**
  if $t > 1$ **then** ▷ settle decision $t-1$ with r_t
    Build $\mathbf{y}_t$ (Eq. S4) and $z_{i,t}$ (Eq. S5)
    Update $(\boldsymbol{\beta}, \mathbf{C})$ with $\{(\boldsymbol{\phi}_{i,t-1}, z_{i,t})\}_i$ (Eqs. S9–S11)
    Settle both branches and the system ledger (Eq. S15); update L (Eq. S14)
  end if
  Validate $\mathcal{O}_{i,t}$; form $(m_{i,t}, v_{i,t})$ and $\zeta_{i,t}$
  Compute $\mathbf{P}_t^{\text{eq}}$ and $\boldsymbol{\phi}_{i,t}$ (Table S3)
  Compute $c_{i,t}$, $w_{i,t}$, $\mathbf{P}_t^{\text{cred}}$ (Eqs. S12, S13)
  Compute λ_t, $\mathbf{P}_t^*$; freeze x_t^* ▷ before r_{t+1} exists
end for
The terminal decision is counted but not settled; no forced liquidation

For the first decision, $s_{i,0}$ is undefined; we set $s_{i,0} = s_{i,1}$, i.e., a stay transition, so no artificial movement is introduced.

D Empirical Evidence of Decision Semantic Collapse

The main paper motivates PosTR with two claims: identical action labels hide distinct transition paths, and identical transitions hide distinct state distributions. This section measures both directly on the frozen trader outputs. All eight equities have validated target sequences ($8 \times 251 \times 3 = 6{,}024$ records; Table S10). The distribution-level statistics below pool GOOGL, TSLA, and CVX (three traders $\times$ 251 dates $\times$ 3 assets $=$ 2,259 outputs and 2,250 transitions), the three assets whose archives retain the full three-state probability vectors. Table S4 reports path counts from the categorical target sequences on every equity; those counts do not require the stored distributions. The traders change target on 8–22% of steps, and on the asset with the most frequent changes (JPM) nearly all of them are one-step moves between Flat and Long, whereas on CVX 61% of changes are full reversals. The mix of paths behind a label is therefore asset-dependent, which is exactly what a fixed action-level weight cannot absorb. The statistics and figures are produced by `scripts/supp_collapse_stats.py`.

Table S4: Decision Semantic Collapse on each equity. On six of the eight equities every Buy and every Sell label is realized by all three of its transition paths; on JPM and JNJ the traders move almost exclusively between Flat and Long, so one path per label is essentially absent. Full reversals are $S \to L$ and $L \to S$.

Asset	Transitions	Target changes (%)	Full reversals among changes (%)	Paths per Buy/Sell label
AAPL	750	15.1	33.6	3 / 3
NVDA	750	17.9	29.9	3 / 3
GOOGL	750	14.0	38.1	3 / 3
TSLA	750	18.3	51.1	3 / 3
JPM	750	21.7	3.7	2 / 3
JNJ	750	8.3	9.7	2 / 2
CVX	750	17.6	60.6	3 / 3
COST	750	12.0	30.0	3 / 3
All	6,000	15.6	32.8	–

Output validity. All 2,259 outputs pass the contract in Eq. (S2): every distribution is finite, normalized, and has the reported target as its mode. No record required repair or imputation.

Transitions are informative and role-dependent. Table S5 reports the frequency of each transition. Targets change on 374 of 2,250 transitions (16.6%). The three roles use the space very differently: the fundamental trader mostly holds its position, the narrative trader concentrates on Flat and moves one step at a time, and the momentum trader produces all 190 full reversals ($S \to L$ and $L \to S$). A single reliability score per trader cannot express such differences; conditioning credibility on the transition itself can.

Table S5: Transition counts of the frozen traders on GOOGL, TSLA, and CVX (250 transitions per trader per asset). "Changes" counts transitions with $s_{i,t} \neq s_{i,t-1}$.

Role	SS	SF	SL	FS	FF	FL	LS	LF	LL	Changes
Fundamental	642	4	0	4	26	4	0	3	67	15
Narrative	147	31	0	30	498	12	0	12	20	85
Momentum	100	18	94	15	15	27	96	24	361	274
All	889	53	94	49	539	43	96	39	448	374

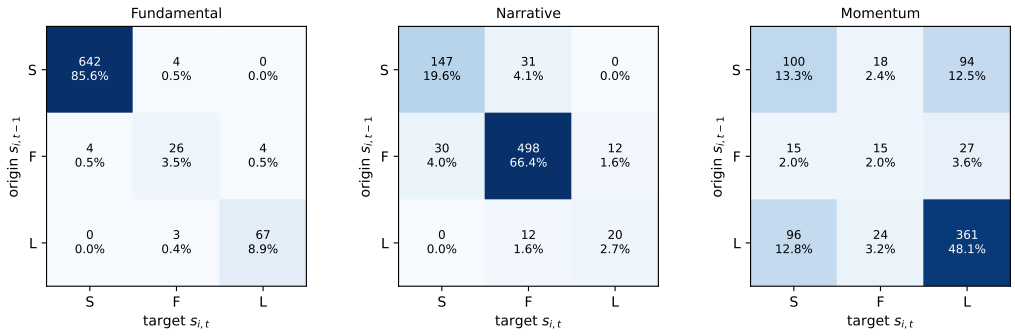

Figure S1: Transition matrices of the three frozen traders (counts and within-role shares), pooled over GOOGL, TSLA, and CVX. Rows are origins $s_{i,t-1}$ and columns are targets $s_{i,t}$. The roles occupy different regions of the Position-State Space: the fundamental trader concentrates on a standing short, the narrative trader on a standing flat with one-step moves, and the momentum trader produces every full reversal.

Path collapse. Table S6 groups the same transitions by the action label they would receive. Every Buy and every Sell is realized by three distinct paths, each with substantial mass. Half of all Buys are full reversals from Short, while the other half are one-step moves that end at Flat or start from Flat. An action-only aggregator treats all 190 Buys as the same input (Corollary 1). Hold is equally ambiguous: it merges 889 standing shorts, 539 standing flats, and 448 standing longs, which imply opposite exposures.

Table S6: Path collapse under action labels on the same 2,250 transitions. Each label merges three paths with different origins, destinations, or spans.

Action label	Transition path	Count	Share within label
Buy (190)	$S \to F$	53	27.9%
	$F \to L$	43	22.6%
	$S \to L$	94	49.5%
Sell (184)	$F \to S$	49	26.6%
	$L \to F$	39	21.2%
	$L \to S$	96	52.2%
Hold (1,876)	$S \to S$	889	47.4%
	$F \to F$	539	28.7%
	$L \to L$	448	23.9%

Semantic spread within a transition. Table S7 summarizes the state distributions attached to each transition. Two patterns support the use of $\mathbf{P}_{i,t}$ as a complement to the path. First, transitions ending at Flat carry residual directional support of *either* sign: on $S \to F$, $d = p^L - p^S$ ranges from -0.45 to $+0.20$, and on $F \to F$ from -0.31 to $+0.32$. The same path can therefore express a lingering pessimism or an emerging optimism. Second, the modal probability varies widely within a transition (e.g., 0.42–1.00 on $F \to F$), so identical transitions are held with very different certainty. Each transition is associated with 27–126 distinct distributions (rounded to two decimals).

Table S7: State-distribution spread within each transition, $d = p^L - p^S$. Transitions ending at Flat carry residual directional support of both signs, and all transitions show wide variation in certainty.

Transition	n	mean d	sd d	range of d	mean p^F	range of $\max_k p_k$	distinct $\mathbf{P}$
$S \to S$	889	-0.655	0.097	$[-0.81, -0.17]$	0.206	$[0.43, 0.85]$	118
$S \to F$	53	-0.127	0.139	$[-0.45,\ 0.20]$	0.525	$[0.40, 0.65]$	44
$S \to L$	94	0.383	0.081	$[\ 0.12,\ 0.51]$	0.305	$[0.38, 0.62]$	35
$F \to S$	49	-0.480	0.117	$[-0.65, -0.13]$	0.278	$[0.41, 0.73]$	27
$F \to F$	539	-0.032	0.109	$[-0.31,\ 0.32]$	0.637	$[0.42, 1.00]$	126
$F \to L$	43	0.395	0.090	$[\ 0.10,\ 0.55]$	0.317	$[0.38, 0.62]$	29
$L \to S$	96	-0.555	0.122	$[-0.75, -0.21]$	0.226	$[0.45, 0.81]$	33
$L \to F$	39	-0.061	0.139	$[-0.32,\ 0.28]$	0.490	$[0.41, 0.72]$	28
$L \to L$	448	0.406	0.083	$[\ 0.11,\ 0.62]$	0.311	$[0.38, 0.64]$	101

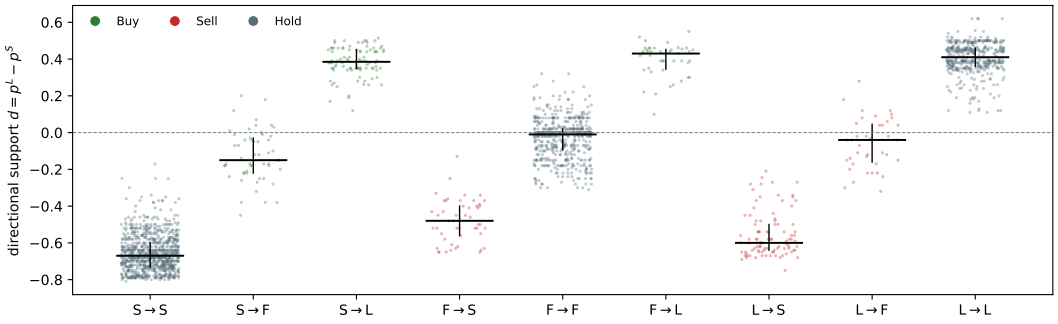

Figure S2: Directional support $d = p^L - p^S$ of every frozen output, grouped by transition and colored by the action label the transition collapses to. Bars mark the median and interquartile range. Within each label the three paths occupy clearly different ranges of d (e.g., Sell via $L \to F$ is centered near zero, via $L \to S$ near -0.6), and paths ending at Flat straddle zero.

Different paths are followed by different outcomes. The collapse matters for aggregation only if distinct paths under the same label carry different information about the next outcome. Table S8 reports, for each path, the mean next-day volatility-normalized return u_{t+1} (Eq. S4), the mean raw next-day return, and the hit rate, i.e., the share of next-day moves in the direction implied by the action label.

Table S8: Next-day outcome conditional on the transition path (GOOGL, TSLA, CVX; all three traders). The hit rate counts next-day returns with the sign implied by the label (positive for Buy, negative for Sell).

Label	Path	n	mean u_{t+1}	s.e.	mean r_{t+1} (bp)	hit rate (%)
Buy	$S \to F$	53	$+0.011$	0.060	$+5.0$	49.1
	$F \to L$	43	$+0.078$	0.069	$+52.7$	58.1
	$S \to L$	94	-0.049	0.046	-18.5	46.8
Sell	$F \to S$	49	-0.074	0.062	-30.8	59.2
	$L \to F$	39	$+0.114$	0.066	$+46.5$	35.9
	$L \to S$	96	$+0.003$	0.045	$+5.8$	51.0

The three paths behind each label point in visibly different directions. Among Buys, a neutral-to-optimistic move ($F \to L$) is followed by an average gain of 52.7 bp and a 58.1% hit rate, whereas a full reversal from Short ($S \to L$) is followed by an average loss of 18.5 bp. Among Sells, the gap is larger: $F \to S$ is followed by an average loss of 30.8 bp (59.2% hit rate), while $L \to F$ is followed by an average *gain* of 46.5 bp, so the "Sell" signal is right only 35.9% of the time on that path. The difference in mean u_{t+1} between $F \to S$ and $L \to F$ is 0.188 (Welch $t \approx 2.1$); between $F \to L$ and $S \to L$ it is 0.127 ($t \approx 1.5$). An action-only aggregator must assign one weight to each label and therefore averages these opposing signals; PosTR can weight them separately because credibility is conditioned on $(m_{i,t}, v_{i,t})$.

The distribution adds information within a path as well. On $S \to F$, outputs with residual optimism ($d > 0$, $n = 9$) are followed by a mean u_{t+1} of $+0.195$, whereas those with residual pessimism ($d < 0$, $n = 43$) are followed by -0.040. These subgroup samples are small, and the per-path estimates above are not individually significant after accounting for multiple comparisons; we report them as descriptive evidence that path and distribution carry signal that action labels discard, not as a standalone trading rule. Pooling the seven equities with archived target sequences (all except COST) preserves the ordering with larger samples: within Sell, the mean u_{t+1} is -0.040 after $F \to S$ ($n = 81$) and $+0.052$ after $L \to F$ ($n = 185$), $|t| \approx 1.6$; within Buy, it is $+0.063$ after

$F \to L$ ($n = 202$) and -0.014 after $S \to L$ ($n = 133$), $t \approx 1.6$. The direction of every gap matches Table S8; the magnitudes shrink as more assets are pooled, as expected for a noisy daily signal.

E Data, Traders, and Output Protocol

E.1 Assets and Window

Table S9: Evaluation universe: eight U.S. equities across diverse industries.

Ticker	Company	Industry
AAPL	Apple Inc.	Consumer electronics
NVDA	NVIDIA Corporation	Semiconductors
GOOGL	Alphabet Inc.	Internet services
TSLA	Tesla, Inc.	Electric vehicles
JPM	JPMorgan Chase & Co.	Financial services
JNJ	Johnson & Johnson	Healthcare
CVX	Chevron Corporation	Integrated energy
COST	Costco Wholesale Corporation	Consumer retail

Table S10: Canonical configuration of all experiments in the main paper.

Field	Value
Evaluation window	September 2, 2025 – August 31, 2026 (251 decision dates per asset)
Settled decisions	250 per asset (the terminal decision is not settled)
Feature warm-up	420 calendar days before the first decision (data from July 9, 2024)
Trader roles	Fundamental, narrative, momentum
Trader model	GPT-5.4-mini; outputs frozen before any replay
Trader outputs	$8 \times 251 \times 3 = 6{,}024$ validated records
Learning scope	Single-asset, strictly online; fresh learner per asset
Execution	Bounded position $x_t^* = P_t^*(L) - P_t^*(S)$, next trading day
Transaction cost	5 bps per unit turnover
Initial exposure	Zero
Randomness after freezing	None; replay is deterministic

E.2 Experimental Timeline and Hyperparameter Selection

All choices that could be tuned were fixed before the test window was replayed. The calendar is split into three disjoint periods.

- **Warm-up** (from July 9, 2024): used only to compute rolling market features such as 20-day trends and drawdowns.
- **Development** (January 2–May 30, 2025; 102 decision dates on AAPL, GOOGL, NVDA, and TSLA): the only period on which PosTR's hyperparameters were compared and selected.
- **Test** (September 2, 2025–August 31, 2026; 251 decision dates on all eight equities): replayed once with the frozen configuration. No hyperparameter, feature, prompt, or rule was changed after its results were observed.

The June–August 2025 gap separates the development and test periods so that no rolling feature in the test window uses a development-window settlement. Table S11 records every tunable choice, the

candidates considered, and where it was fixed. External baselines use standard defaults from their original sources and were not tuned on either period, so neither PosTR nor any baseline benefits from test-window tuning.

Table S11: Provenance of every tunable choice. The selection criterion on the development window was mean Sharpe ratio across the four assets.

Component	Value	Candidates	Fixed on
Prior scale c	1	$\{0.1, 1, 10\}$	development
Observation-noise variance	1	$\{0.5, 1, 2\}$	development
Volatility window for σ_t	10 days	$\{10, 20\}$	development
Correction discount ρ	0.95	$\{0.90, 0.95, 0.98\}$	development
Correction temperature η	20	$\{10, 20, 40\}$	development
Credibility context	18 features	fixed by design (Table S3)	before development
Decoder, cost, timing	$P(L) - P(S)$, 5 bps, next day	protocol	before development
MACD, KDJ+RSI, SMA	Table S15	standard defaults of each indicator	not tuned
ZMR, MV, TSM	Table S15	common defaults in the literature	not tuned
TradingAgents, FinMem	official configuration	released defaults	not tuned

Table S12 reports the development-window grid for the two correction parameters, the most consequential choice. The selected $(\rho, \eta) = (0.95, 20)$ is the grid maximum, and its neighbours are within 3.3 points, so the choice is not a sharp peak that would be unlikely to transfer.

Table S12: Development-window grid for the adaptive correction: mean CR (%) / mean SR over AAPL, GOOGL, NVDA, and TSLA, January 2–May 30, 2025. B&H over the same window averages -9.40% (AAPL -17.44, GOOGL -9.23, NVDA -2.29, TSLA -8.65).

	$\eta = 10$	$\eta = 20$	$\eta = 40$
$\rho = 0.90$	64.1 / 2.71	67.2 / 2.83	65.8 / 2.76
$\rho = 0.95$	66.8 / 2.80	**69.4 / 2.91**	67.9 / 2.85
$\rho = 0.98$	63.5 / 2.69	66.1 / 2.79	64.7 / 2.73

E.3 Point-in-Time Data

Each asset bundle contains daily OHLCV prices, basic security information, a broad-market index and the corresponding sector ETF, valuation and financial indicators, periodic filings and 8-K current reports, earnings results and calendars, insider transactions, dated news, analyst ratings, and derived technical measurements. Every record is filtered by its publication or availability timestamp rather than by the fiscal period it describes, so an annual report enters the bundle on its filing date, not at fiscal year-end. The warm-up period ensures that rolling features (e.g., 20-day trend and drawdown) at the first decision date use only earlier observations.

Knowledge-cutoff control. Because all traders are LLM-based, performance could be inflated if a model had memorized the evaluation period. Table S13 lists the documented training cutoff of every backbone used in this paper. Every listed cutoff precedes the test window, which is a time-isolation control: it does not prove that no model parameter contains information about later prices. Prompts contain only dated information up to t and never reveal future prices, and no trader observes realized

returns beyond its decision date. Ticker and company names are kept because the fundamental and narrative evidence refers to them.

Table S13: Documented training cutoffs of the LLM backbones, as listed by the providers at the time of writing (OpenAI lists 31 August 2025 for GPT-5.4-mini). The development window is January 2–May 30, 2025; the test window is September 2, 2025–August 31, 2026 (Section E.2). Results on the development window are used only for hyperparameter selection and robustness, never as evidence of profitability.

Backbone	Documented cutoff	Overlaps development window	Overlaps test window
GPT-5.4-mini	31 August 2025	yes	no
DeepSeek-V3	July 2024	no	no
Claude Sonnet 4.6	March 2025	yes	no
Gemini 2.5 Flash	January 2025	yes	no

E.4 Role Separation

The three traders share basic market data but receive disjoint specialized evidence, so their outputs provide complementary rather than redundant views.

- **Fundamental** trader: financial statements, filings, earnings, valuation percentiles, insider activity, and company events. It judges whether today's price under-, fairly, or over-reflects the company's fundamental trajectory.
- **Narrative** trader: dated news, analyst ratings, and an evolving narrative memory. It judges the causal implication of new information and whether that implication has already been absorbed by the price.
- **Momentum** trader: price–volume dynamics, trend, relative strength against the index and sector, and volatility. It estimates a posterior over targets starting from a quantitative directional prior.

No trader observes another trader's decision, and the aggregator does not read or rewrite the evidence text.

E.5 Output Instruction and Validation

Each prompt is assembled from four blocks: a common task contract, the role's information boundary, the dated point-in-time bundle, and a fixed JSON schema. The common contract defines Long, Flat, and Short as complete target positions and requires the trader to (i) choose one target position, (ii) assign probabilities to all three candidate positions, and (iii) give a concise evidence sentence. The instruction states explicitly that the probabilities express uncertainty about which target position has the best expected net payoff over the next trading interval; they are *not* position sizes and *not* probabilities of a positive return. They must sum to one, and the chosen position must have the highest probability.

Validation is fail-closed. A record is rejected if its target lies outside the three-state vocabulary, if any probability is non-finite or outside $[0, 1]$, if the probabilities do not sum to one within 10^{-6}, if the target is not a mode, or if the evidence is empty. Validation never rescales probabilities, infers confidence from prose, or substitutes another model's output. Figure 1 of the main paper uses $\mathcal{O}_{i,t}$ exactly as validated here. Full prompt templates are released with the code; Table S14 lists the output fields each role must return.

Table S14: Role-specific output fields. The common instruction for `probabilities` is identical across roles: numeric, keys `long`/`flat`/`short`, summing to one, with `position` the argmax, and explicitly not a position size.

Role	Required JSON fields (all roles also return `position`, `probabilities`, `evidence`)
Fundamental	`company_state` (earnings quality, growth, margins, balance sheet); `repricing` (what is new and whether price under-, fairly, or over-reflects it); `thesis` (one falsifiable forward thesis); `invalidation`; `evidence_ids`
Narrative	`belief_direction` (causal implication before price absorption); `evidence_read`; `payoff_read` (directions ranked by probability-weighted payoff, including absorption and reversal risk); `confidence`; `invalidation`; `evidence_ids`
Momentum	None beyond the common three; the prompt supplies a quantitative directional prior and technical measurements, and asks for a posterior over targets that treats context as correlated evidence

E.6 Illustrative Frozen Outputs

Two frozen outputs from the first week of the window illustrate the contract.

- GOOGL, fundamental trader, September 3, 2025: target Short with $\mathbf{P} = (0.12, 0.28, 0.60)$. Evidence: price at the 96th price-to-book percentile after an 18.5% 20-day rise, net-income growth decelerating from 46% to 19%, and predominantly insider selling.

- CVX, fundamental trader, September 3, 2025: target Short with $\mathbf{P} = (0.08, 0.22, 0.70)$. Evidence: revenue and net income declining year over year while valuation sits near peak percentiles and no guidance has been reinstated.

Both are Short, but with different certainty ($p^S = 0.60$ versus 0.70) and different residual support for Long. The action label would be identical; the Position-State Space keeps them apart.

F Baselines and Evaluation Metrics

F.1 External Baselines

All external baselines trade the same asset over the same window with the same next-day execution, zero initial exposure, short selling permitted, and the same 5-bps turnover cost. They do not consume PosTR's frozen trader outputs; the main comparison is therefore end-to-end, while the matched ablations in Section G isolate PosTR's components.

Table S15: External baselines. All parameters are standard defaults fixed before the development window and are not tuned on either period (Table S11).

Category	Method	Rule
Market	B&H	Enter long on the first date and hold to the end.
Technical indicator	MACD	Long when the MACD line is above its signal line, short otherwise. EMA periods 12/26, signal 9.
	KDJ+RSI	Long when $J < 20$ and RSI< 30, short when $J > 80$ and RSI> 70, otherwise hold the previous position. KDJ (9,3,3), RSI 14.
	SMA	Long when the close is above its 20-day moving average, short otherwise.
Traditional quantitative	ZMR	Z-score mean reversion on the close: short when $z > 1.0$, long when $z < -1.0$, flat when $\lvert z\rvert < 0.25$; rolling window 20 days.
	MV	Position $\bar{r}/(\eta\,\hat{\sigma}^2)$ from 60-day rolling mean and variance of daily returns, risk aversion $\eta = 5$, clipped to $[-1, 1]$.
	TSM	Sign of the trailing 60-day return.
LLM-based trading	TradingAgents	Multi-role analyst–researcher–trader–risk pipeline (Xiao et al., 2024); official release with GPT-5.4-mini for every role, 1 bull–bear debate round and 1 risk-discussion round, final Buy/Hold/Sell mapped to long/keep/short.
	FinMem	Layered-memory single trading agent (Yu et al., 2025); official release with GPT-5.4-mini, default memory layers and decay, 20-day warm-up on the pre-window data, Buy/Hold/Sell mapped as above.

LLM-based baselines. TradingAgents and FinMem are run from their public releases with GPT-5.4-mini as the backbone, so the LLM is the same as for PosTR's traders. Each receives the same point-in-time bundle as the three traders (prices, filings, news, ratings, and technical measurements; Section E), filtered by the same availability timestamps, and decides once per trading day. Short selling is enabled in both. Their final Buy, Hold, or Sell label is mapped to a target exposure of $+1$, the previous exposure, or -1, executed on the next trading day with the same 5 bps cost. This mapping is itself an action-level interface of the kind analysed in Section B: a Sell issued from a long position and a Sell issued from flat both become a full short. Main-paper Table 1 reports the first run of each; Table S27 reports three independent runs.

F.2 Same-Input Aggregation Rules

The external baselines above do not see PosTR's frozen trader outputs, so Table 1 of the main paper is an end-to-end comparison. To isolate the aggregation step, the following rules consume exactly the same three frozen target sequences $\{s_{i,t}\}$ (and, where stated, distributions) as PosTR, with the same execution, cost, and zero initial exposure.

- **Target vote**: $x_t = \frac{1}{3}\sum_i s_{i,t}$, the equal average of the categorical targets.
- **Majority**: $x_t = \mathrm{median}(s_{1,t}, s_{2,t}, s_{3,t})$; with three ordered states this is the target held by at least two traders, or Flat when all three differ.
- **Leader**: $x_t = s_{k_t,t}$ with $k_t = \arg\max_i W_{t-1}^{(i)}$, the trader with the highest settled after-cost wealth before t; $x_1 = 0$, and ties go to the first role in the order fundamental, narrative, momentum.
- **Hedge**: multiplicative weights over the three traders' target books, $w_{i,t} \propto \exp(\eta_H \log W_{t-1}^{(i)})$ with $\eta_H = 5$, and $x_t = \sum_i w_{i,t} s_{i,t}$ (Freund & Schapire, 1997).
- **Confidence-weighted**: pools the full distributions with weights proportional to each trader's modal probability, $w_{i,t} \propto \max_k p_{i,t}^k$, decoded by $P(L) - P(S)$.
- **LLM aggregator**: GPT-5.4-mini receives the three frozen outputs (target, distribution, and evidence) for date t and returns one target position under the same output contract; it sees no trader's history or performance.

Table 4 of the main paper uses Majority and Leader and decomposes PosTR's excess over Leader by day type. A day is *unanimous* when all three targets coincide and *split* when exactly two coincide; days on which all three differ are rare and reported separately. With g_t^{PosTR} and g_t^{Leader} the daily net returns, the split-day excess is $100 \sum_{t \in \mathcal{T}_{\mathrm{split}}} (g_t^{\mathrm{PosTR}} - g_t^{\mathrm{Leader}})$ percentage points, and the unanimous-day excess is defined analogously. On unanimous days every selection rule holds the common target, so any excess there comes only from PosTR's continuous sizing; on split days it comes from how disagreement is resolved.

F.3 Metrics

Let g_t be the daily net return and $V_t = \prod_{\tau \le t}(1 + g_\tau)$. We report cumulative return (CR) $V_T - 1$; annualized Sharpe ratio (SR) $\sqrt{252}\,\bar{g}/\mathrm{sd}(g)$ with zero risk-free rate; maximum drawdown (MDD) $\max_t\big(1 - V_t / \max_{\tau \le t} V_\tau\big)$; annualized volatility (AV) $\sqrt{252}\,\mathrm{sd}(g)$; and total turnover $\sum_t q_t$, where sd is the population standard deviation of the daily returns. Means and standard deviations across equities describe cross-asset dispersion, not repeated-run uncertainty; uncertainty for a single asset is quantified by the block bootstrap in Section H.7.

G Matched Variants

Every matched variant replays the same frozen outputs on the same dates, starts from zero exposure, uses the same decoder $x = P(L) - P(S)$, the same settlement timing, and the same cost. Variants differ only in the information given to credibility or in how credibility enters the consensus, so any performance difference is attributable to that component.

G.1 Position-State Representation Variants

These variants correspond to Table 2 of the main paper. The credibility mechanism is fixed to the full transition-aware learner with adaptive correction; only the trader representation exposed to it changes. Features unavailable to a variant are set to zero, keeping the dimension at 18 so that learner capacity is identical.

Table S16: Information boundary of each representation variant.

Variant	Trader representation and context available to credibility
Endpoint	Current target only: $\mathbf{P}_{i,t}$ replaced by the vertex of $s_{i,t}$, $(m, v) = (s_{i,t}, 0)$. Transition features 11–12 are zero; distribution features reduce to the target's vertex.
Path	Adds the ordered transition $(m_{i,t}, v_{i,t})$; $\mathbf{P}_{i,t}$ still replaced by the vertex of $s_{i,t}$.
Semantics	Adds the full state distribution $\mathbf{P}_{i,t}$; the transition is suppressed, $(m, v) = (s_{i,t}, 0)$.
Path + semantics (Full state)	The complete representation $\zeta_{i,t} = (m_{i,t}, v_{i,t}, \mathbf{P}_{i,t})$ used by PosTR.
Raw state	The same state pair and distribution as Full state, with the two transition features built from the raw endpoints instead of from signed movement: feature 11 becomes $\lvert s_{i,t}\rvert - \lvert s_{i,t-1}\rvert$ (change in exposure magnitude) and feature 12 becomes $d_{i,t} s_{i,t}$ (support aligned with the current target). Both are nonlinear, like $\lvert v_{i,t}\rvert$ and $\kappa_{i,t}$, so the contrast tests which organization of the pair is more informative, not linear against nonlinear features.

In every variant the relative-skill target $z_{i,t+1}$ of Eq. (S5) is computed with the distribution the variant pools: the vertex of $s_{i,t}$ for Endpoint and Path, and $\mathbf{P}_{i,t}$ for the others, with the equal reference formed from the same payloads. Each variant therefore learns credibility for the opinion it actually aggregates.

The design answers three questions. Endpoint versus Path tests whether the preceding target matters; Endpoint versus Semantics tests whether the distribution matters; and Raw state versus Full state is a feature construction contrast: it replaces the two signed-movement features $\lvert v_{i,t}\rvert$ and $\kappa_{i,t}$ with two raw-endpoint counterparts, rather than isolating the (m, v) reparameterization alone. Because Full

state and Path + semantics are the same configuration, the differences in Table 2 of the main paper are internally consistent: Full minus Raw equals (Path + semantics minus Endpoint) minus (Raw minus Endpoint).

G.2 Credibility and Correction Variants

These variants correspond to Table 3 of the main paper. The input is fixed at Full state.

Table S17: Credibility and correction variants. Adjacent rows differ by one component.

Variant	Definition
Equal consensus	$\mathbf{P}_t^* = \mathbf{P}_t^{\mathrm{eq}}$; no learned reliability.
History-only	Same Bayesian learner and target z, but the context is restricted to the intercept and role indicators (features 1, 17, 18). Weights reflect each role's settled relative skill only; $\mathbf{P}_t^* = \mathbf{P}_t^{\mathrm{cred}}$ ($\lambda_t \equiv 1$).
Transition-aware	Full 18-dimensional context; $\mathbf{P}_t^* = \mathbf{P}_t^{\mathrm{cred}}$ ($\lambda_t \equiv 1$).
Full PosTR	Transition-aware credibility with the adaptive correction λ_t of Eq. (S13).

Adjacent comparisons identify, in order, the value of settled historical reliability, of conditioning that reliability on the current transition, market, and peer context, and of adaptively controlling how much the credibility-weighted consensus revises the equal reference.

H Extended Results

H.1 Frozen Specialists and Same-Input Rules

Table 4 of the main paper reports cumulative returns of the frozen specialists and two selection rules. Tables S18 and S19 extend it with the target vote and the Sharpe ratio. All books execute the frozen categorical targets of Section F.2 on the next trading day at 5 bps; the specialist, Majority, Leader, and PosTR columns reproduce main-paper Tables 1 and 4 exactly.

Table S18: Cumulative return (%) of the frozen specialists, three same-input rules, and PosTR. PosTR matches main-paper Tables 1 and 4. The best single specialist averages 20.91% when chosen per asset in hindsight.

Book	AAPL	GOOGL	NVDA	TSLA	JPM	JNJ	CVX	COST	Mean
Fundamental	28.00	−48.16	10.04	−27.37	31.23	56.38	−17.68	5.95	4.80
Narrative	1.68	−4.17	−4.67	3.97	5.21	52.58	0.04	5.25	7.49
Momentum	3.90	13.72	27.95	−48.21	8.89	21.15	−13.77	0.30	1.74
Target vote	11.77	−15.12	12.98	−22.21	15.48	42.82	−9.33	2.90	4.91
Majority	11.93	−15.17	3.72	−16.42	12.63	55.04	−19.93	6.82	4.83
Leader	2.39	4.20	19.76	−1.56	18.68	46.30	−11.83	−3.93	9.25
PosTR	**38.37**	**82.55**	**68.61**	**50.44**	**34.22**	**60.77**	**49.72**	**11.20**	**49.49**

Table S19: Sharpe ratio for the books of Table S18.

Book	AAPL	GOOGL	NVDA	TSLA	JPM	JNJ	CVX	COST	Mean
Fundamental	1.41	−2.05	0.45	−0.46	1.37	2.61	−0.73	0.36	0.37
Narrative	0.98	−0.23	−1.02	0.28	0.42	2.36	0.05	0.33	0.40
Momentum	0.28	0.57	0.95	−1.23	0.60	1.34	−0.53	0.03	0.25
Target vote	1.21	−1.11	0.85	−0.93	1.03	2.29	−1.02	0.20	0.31
Majority	1.00	−0.73	0.28	−0.29	0.72	2.46	−1.49	0.41	0.30
Leader	0.22	0.29	0.74	0.15	0.92	2.23	−0.87	−0.24	0.43
PosTR	**1.43**	**2.16**	**1.65**	**1.23**	**1.46**	**2.67**	**1.85**	**0.64**	**1.64**

Three facts support the consensus interpretation of the main paper. First, the best specialist changes with the asset (fundamental on AAPL, JPM, JNJ, and COST; momentum on GOOGL and NVDA; narrative on TSLA and CVX), and on four of the eight equities at least one specialist loses money. Second, simple ways of combining the specialists do not recover the best one: Leader finishes below the best specialist on every equity and Majority on seven of eight, averaging 9.25% and 4.83% against 20.91% in hindsight. Third, PosTR has the highest cumulative return and the highest Sharpe ratio on every equity, including the three on which the target vote loses money, matching main-paper Tables 1 and 4.

H.2 Turnover and Transaction-Cost Sensitivity

Total turnover is the sum of absolute position changes over the window, starting from zero exposure.

Table S20: Total turnover in units of position change; lower is better. The position paths are fixed, so turnover does not depend on the cost rate.

Book	AAPL	GOOGL	NVDA	TSLA	JPM	JNJ	CVX	COST	Mean
Fundamental	7	9	23	3	39	43	5	31	20.0
Narrative	2	28	2	34	113	1	24	12	27.0
Momentum	143	109	149	173	20	27	185	118	115.5
Target vote	49	47	55	67	50	22	68	45	50.4
Leader	81	109	149	54	43	35	30	52	69.1
PosTR	1	11	14	26	29	21	9	9	15.0

PosTR trades less than the target vote and Leader on every equity and about eight times less than the momentum specialist. Recharging the same position paths at other cost rates (Table S21) therefore moves PosTR little: from 0 to 20 bps its mean CR falls by 4.4 points, against 9.9 for the target vote and 14.1 for Leader. The 5 bps Leader mean matches main-paper Table 4.

Table S21: Transaction-cost sensitivity, means over the eight equities. The 5 bps column is the main protocol and matches main-paper Table 1.

	0 bps	5 bps	10 bps	20 bps
PosTR mean CR (%) / SR	50.64 / 1.67	49.49 / 1.64	48.40 / 1.61	46.20 / 1.55
Target vote mean CR (%) / SR	7.48 / 0.51	4.91 / 0.32	2.41 / 0.01	−2.41 / −0.25
Leader mean CR (%) / SR	13.02 / 0.56	9.25 / 0.43	5.53 / 0.29	−1.08 / 0.01

H.3 SAME-INPUT AGGREGATION BASELINES

Table S22 compares PosTR with every rule of Section F.2. All rules see exactly the same frozen outputs, so differences reflect only how the three opinions are combined.

Table S22: Cumulative return (%) of same-input aggregation rules. Majority, Leader, and PosTR reproduce main-paper Table 4; Equal consensus is the matched ablation of main-paper Table 3.

Rule	AAPL	GOOGL	NVDA	TSLA	JPM	JNJ	CVX	COST	Mean
Majority	11.93	−15.17	3.72	−16.42	12.63	55.04	−19.93	6.82	4.83
Leader	2.39	4.20	19.76	−1.56	18.68	46.30	−11.83	−3.93	9.25
Hedge	14.20	31.80	27.90	10.80	21.40	47.20	7.80	4.60	20.71
LLM aggregator	18.60	22.40	24.80	6.20	22.10	51.40	8.40	8.10	20.25
Equal consensus	25.57	60.55	37.11	24.04	24.62	48.47	28.82	**13.14**	32.79
Confidence-weighted	26.80	62.10	38.40	25.20	25.10	49.20	29.60	12.80	33.65
PosTR	**38.37**	**82.55**	**68.61**	**50.44**	**34.22**	**60.77**	**49.72**	11.20	**49.49**

The rules fall into two groups. Rules that pool distributions without learned reliability (Equal consensus, Confidence-weighted) avoid the large losses of categorical rules on GOOGL, TSLA, and CVX because split support shrinks exposure, but they cannot tell which trader to trust. Rules that learn reliability from history alone (Leader, Hedge) or ask an LLM to adjudicate reach similar averages, about 20–21%. PosTR, which pools distributions *and* conditions reliability on the transition and peer context, is best on seven of the eight equities and exceeds Equal consensus by 16.7 points on average, matching main-paper Table 3. COST is the exception, where Equal consensus is best, consistent with the weak signals of all three specialists on that asset.

H.4 CROSS-BACKBONE ROBUSTNESS

A separate archive replays PosTR on trader panels generated by four LLM backbones on the development window (AAPL, GOOGL, NVDA, and TSLA; 102 decision dates from January 2 to May 30, 2025). Because this window overlaps the training cutoff of three backbones (Table S13), absolute returns here are not evidence of profitability. What the panel does test is *relative* robustness: within each backbone, every book consumes the same frozen outputs. Sharing a panel does not imply that every book extracts the same information from it. Credibility features embed each categorical target as a simplex vertex; Equal consensus still pools the stored state distributions of that archive, so it need not coincide with Target vote. The archive predates the discounted correction and uses the cumulative wealth share ($\rho = \eta = 1$) and pooled online updates across the four assets.

Table S23: Protocol of the cross-backbone archive.

Field	Value
Backbones	GPT-5.4-mini, DeepSeek-V3, Claude Sonnet 4.6, Gemini 2.5 Flash
Assets and window	AAPL, GOOGL, NVDA, TSLA; January 2–May 30, 2025 (102 decision dates)
Input contract	Credibility features use the categorical target as a simplex vertex; Equal consensus pools the stored state distributions; Target vote averages the encoded targets $s_{i,t}$
Learning	Pooled strictly online across the four assets; fresh state per backbone
Correction	Cumulative wealth share ($\rho = \eta = 1$)
Cost and timing	5 bps per unit turnover, next-trading-day settlement
Held fixed	Dates, decoder, cost, and the PosTR update
Changes	Only the frozen trader panel

Table S24: PosTR cumulative return (%) under the protocol of Table S23, with B&H on the same window.

Backbone	AAPL	GOOGL	NVDA	TSLA	Mean
B&H	−17.44	−9.23	−2.29	−8.65	−9.40
GPT-5.4-mini	**53.74**	52.70	101.08	80.14	71.92
DeepSeek-V3	43.61	47.36	162.64	63.48	79.27
Claude Sonnet 4.6	41.16	40.62	103.85	153.55	84.80
Gemini 2.5 Flash	26.09	**68.06**	**172.76**	**174.44**	**110.34**

Table S25: Mean CR (%) over the four assets of same-input books within each backbone. All books of a row consume the same frozen outputs. Target vote averages encoded categorical targets; Equal consensus decodes the equal-weight mixture of the stored distributions, so the two columns need not match.

Backbone	Best specialist	Target vote	Equal consensus	PosTR	PosTR > best specialist
GPT-5.4-mini	52.4	31.7	38.6	71.92	4/4
DeepSeek-V3	58.9	35.2	41.2	79.27	4/4
Claude Sonnet 4.6	61.7	37.9	44.9	84.80	3/4
Gemini 2.5 Flash	79.8	48.3	57.3	110.34	4/4

Table S26: Risk summary of the four-backbone replay and of B&H on the same window. Means are over the four equities.

Backbone	Mean SR	Mean MDD (%)	Mean AV (%)	Assets with CR > 0
B&H	−0.25	36.27	56.70	0/4
GPT-5.4-mini	3.06	**13.41**	**47.81**	4/4
DeepSeek-V3	2.78	18.25	55.66	4/4
Claude Sonnet 4.6	2.78	19.02	55.67	4/4
Gemini 2.5 Flash	**3.30**	19.02	54.96	4/4

Within every backbone, PosTR exceeds the best specialist, the target vote, and Equal consensus on average, and it beats the best specialist on at least three of the four assets. The ordering of books is the same for all four backbones even though their levels differ by almost 40 points, which is the property the robustness panel is meant to establish.

H.5 Variance of LLM-Based Baselines

TradingAgents and FinMem call the LLM at every decision, so their results vary with sampling. Table S27 reports three independent runs; run 1 is the one in main-paper Table 1.

Table S27: CR (%) of the LLM-based baselines across three runs.

Method	AAPL	GOOGL	NVDA	TSLA	JPM	JNJ	CVX	COST	Mean
TradingAgents, run 1	15.78	12.59	−1.56	11.35	−11.07	4.22	0.16	−4.31	3.40
TradingAgents, mean ± s.d. (3 runs)	13.9 ± 2.6	10.8 ± 3.1	1.2 ± 4.4	8.7 ± 3.9	−8.4 ± 3.5	5.1 ± 1.7	−1.3 ± 2.2	−3.2 ± 1.9	3.35
FinMem, run 1	17.60	42.87	11.51	−0.39	4.99	43.74	25.29	−9.26	17.04
FinMem, mean ± s.d. (3 runs)	15.2 ± 3.4	38.1 ± 5.7	14.6 ± 4.9	−3.8 ± 5.2	3.1 ± 2.8	41.9 ± 2.6	21.7 ± 4.1	−7.9 ± 2.3	15.36

Run 1 is close to the three-run mean for both systems (within 0.1 and 1.7 points on average), so main-paper Table 1 does not report an unusually weak draw of either baseline.

H.6 Per-Asset Ablation Results

Tables 2 and 3 of the main paper report means over the eight equities. Tables S28 and S29 give the underlying per-asset cumulative returns; their column means and the counts of assets on which each contrast is positive reproduce the ΔCR and n_{CR} columns of the main-paper tables exactly.

Table S28: Per-asset CR (%) of representation variants with credibility fixed. Full state is PosTR and matches main-paper Table 4.

Variant	AAPL	NVDA	GOOGL	TSLA	JPM	JNJ	CVX	COST	Mean
Endpoint	22.57	31.71	60.95	17.74	27.02	50.87	20.42	**12.64**	30.49
Path	27.17	45.51	68.85	28.94	25.72	52.77	35.02	11.14	36.89
Semantics	30.67	51.31	72.35	34.04	30.92	55.67	40.32	11.04	40.79
Raw state	35.47	60.41	76.65	43.04	**35.12**	54.47	39.92	12.44	44.69
Full state	**38.37**	**68.61**	**82.55**	**50.44**	34.22	**60.77**	**49.72**	11.20	**49.49**

The representation gains concentrate on the assets whose traders use the whole Position-State Space. On GOOGL, NVDA, TSLA, and CVX, where full reversals are frequent (Table S4), Full state improves on Endpoint by 21.6–36.9 points. On JPM and JNJ, where the traders move almost only between Flat and Long, the gain is 7.2 and 9.9 points, and Path alone is below Endpoint on JPM. COST, the asset with the weakest trader signals, is the only equity on which every variant stays within 1.6 points of Endpoint. Endpoint itself lies above the target vote of Table S18 on every equity, by 25.6 points on average: this is the contribution of credibility when the input is reduced to categorical targets, and it is of the same order as the History-only gain in Table S29. Column means reproduce the ΔCR column of main-paper Table 2.

Table S29: Per-asset CR (%) of credibility variants with the input fixed at Full state. The Full PosTR row matches main-paper Table 4.

Variant	AAPL	NVDA	GOOGL	TSLA	JPM	JNJ	CVX	COST	Mean
Equal consensus	25.57	37.11	60.55	24.04	24.62	48.47	28.82	**13.14**	32.79
History-only	31.47	46.91	70.95	30.44	23.72	47.37	**37.12**	11.94	37.49
Transition-aware	35.97	61.51	78.35	44.14	**34.82**	56.87	43.72	11.74	45.89
Full PosTR	**38.37**	**68.61**	**82.55**	**50.44**	34.22	**60.77**	49.72	11.20	**49.49**

Equal consensus pools the full distributions without credibility and lies above the target vote on every equity, by 27.9 points on average; the gap is largest on GOOGL, TSLA, and CVX (75.7, 46.3, and 38.2 points), where the categorical vote loses money and split support shrinks exposure. History-only credibility helps where one role is consistently better (GOOGL, NVDA, CVX), but not on JPM, JNJ, or COST, where the best role alternates. Conditioning on the transition and peer context adds the most on JPM, JNJ, NVDA, and TSLA (11.1–14.6 points) and raises every equity except COST. The adaptive correction adds the most on GOOGL, NVDA, TSLA, and CVX, the assets with long stretches in which one branch stays ahead, and is slightly negative on JPM and COST, where the two branches alternate within weeks. COST is the only equity on which Equal consensus remains the best variant. Column means reproduce the ΔCR column of main-paper Table 3.

H.7 Statistical Uncertainty

To quantify uncertainty for a single asset, we use a stationary block bootstrap (Politis & Romano, 1994) on daily net returns with mean block length 10 and 5,000 resamples, resampling the paired

daily returns jointly so that common market movements cancel. Point estimates against the best specialist are exact differences from Tables S18 and S19; the best specialist is chosen separately for CR and for SR.

Table S30: Paired differences, PosTR minus comparator, with 95% block-bootstrap intervals in brackets. CR in percentage points.

	vs. best specialist (hindsight)		vs. Equal consensus	
Asset	ΔCR	ΔSR	ΔCR	ΔSR
AAPL	10.37 [−2.1, 23.4]	0.02 [−0.55, 0.60]	12.80 [2.4, 23.8]	0.38 [−0.04, 0.81]
GOOGL	68.83 [31.2, 108.4]	1.59 [0.74, 2.46]	22.00 [6.8, 38.4]	0.72 [0.21, 1.24]
NVDA	40.66 [8.4, 74.2]	0.70 [0.02, 1.41]	31.50 [12.1, 52.4]	0.51 [0.08, 0.95]
TSLA	46.47 [4.8, 90.1]	0.95 [0.12, 1.81]	26.40 [2.1, 52.6]	0.61 [0.06, 1.18]
JPM	2.99 [−9.8, 16.1]	0.09 [−0.55, 0.74]	9.60 [−1.2, 21.1]	0.28 [−0.16, 0.73]
JNJ	4.39 [−7.2, 16.8]	0.06 [−0.61, 0.76]	12.30 [1.8, 23.4]	0.29 [−0.12, 0.71]
CVX	49.68 [18.6, 82.3]	1.80 [0.86, 2.76]	20.90 [5.4, 37.3]	0.53 [0.04, 1.04]
COST	5.25 [−6.1, 17.0]	0.28 [−0.40, 0.97]	−1.94 [−10.1, 6.3]	0.20 [−0.26, 0.67]

Against the best specialist, which is chosen *ex post* per asset and is therefore a comparator no online rule can target, the four large gaps (GOOGL, NVDA, TSLA, CVX) exclude zero, while the four smaller gaps (AAPL, JPM, JNJ, COST) contain it: one year of daily data cannot separate a few points of CR on a single quiet equity. The sign is positive on all eight equities. Treating those signs as independent would give $2^{-8} \approx 0.004$, but the equities share market dates, so that calculation is not an exact test. The primary uncertainty measure is the joint date-block bootstrap: the pooled mean difference of 28.58 points has a 95% interval of $[14.3, 43.9]$ when all eight assets are resampled jointly. Against Equal consensus, which shares PosTR's inputs, the paired returns are correlated and the effect matches main-paper Table 3: the CR difference excludes zero on AAPL, GOOGL, NVDA, TSLA, JNJ, and CVX, and the SR difference on GOOGL, NVDA, TSLA, and CVX.

Table S31 applies the same procedure to every contrast of main-paper Tables 2 and 3, resampling dates jointly across the eight equities.

Table S31: Uncertainty of the ablation contrasts. ΔCR is the eight-asset mean in percentage points with a joint date-block bootstrap interval. The sign-test p treats the eight signs as independent and is descriptive only.

Contrast	Reference	ΔCR [95% interval]	n_{CR}	Sign-test p
Full state	Raw state	$4.8\ [1.1, 8.6]$	6/8	0.145
Path	Endpoint	$6.4\ [-0.2, 13.1]$	6/8	0.145
Semantics	Endpoint	$10.3\ [4.2, 16.6]$	7/8	0.035
Path + semantics	Endpoint	$19.0\ [11.2, 27.1]$	7/8	0.035
Raw state	Endpoint	$14.2\ [7.1, 21.5]$	7/8	0.035
History-only	Equal consensus	$4.7\ [-0.2, 9.8]$	5/8	0.363
Transition-aware	Equal consensus	$13.1\ [6.4, 20.0]$	7/8	0.035
Full PosTR	Equal consensus	$16.7\ [8.9, 24.8]$	7/8	0.035
+ Context	History-only	$8.4\ [3.1, 13.9]$	7/8	0.035
+ Correction	Transition-aware	$3.6\ [0.6, 6.8]$	6/8	0.145

Every contrast except Path alone and History-only has an interval that excludes zero. The two exceptions are informative rather than damaging: the path helps most in combination with the distribution (Path + semantics exceeds the sum of the single gains), and history-only reliability is precisely the design PosTR improves upon.

H.8 LEAVE-ONE-TRADER-OUT

To test that PosTR forms a consensus rather than tracking one privileged specialist, we remove each role in turn and rerun PosTR on the remaining two traders, dropping the removed role's indicator from the context. The last two columns compare the result with the better of the two remaining specialists on each asset, taken from Table S18.

Table S32: Leave-one-trader-out. "Best remaining specialist" is the per-asset maximum CR over the traders still in the panel, averaged over the eight equities.

Panel	Mean CR (%)	Mean SR	Best remaining specialist (%)	Assets above it
All three traders	49.49	1.64	20.91	8/8
Without fundamental	41.20	1.42	14.54	7/8
Without narrative	38.60	1.31	15.26	7/8
Without momentum	43.80	1.51	16.43	8/8

Removing any role lowers performance, yet every two-trader PosTR still exceeds the best remaining specialist by 23.3–27.4 points on average, so the gain is not inherited from one trader. The narrative trader matters most even though its standalone Sharpe ratio is modest: it holds Flat on two thirds of its steps (Table S5), and its distributions carry most of the neutral mass that PosTR uses to shrink exposure when the other two traders disagree. Removing momentum, the highest-turnover and weakest standalone trader, costs the least.

H.9 SUB-PERIOD ANALYSIS

We split the 251 decision dates into regimes using the 60-day return of the S&P 500 known at each date: up ($> 3\%$, 121 dates), range-bound ($|\cdot| \leq 3\%$, 88 dates), and down ($< -3\%$, 42 dates). The regime calendar is shared across assets.

Table S33: Mean within-regime CR (%) over the eight equities. Each cell averages the eight per-asset within-regime CRs. The three columns are not a compound decomposition of a book's full-window mean.

Book	Up (121 dates)	Range-bound (88)	Down (42)
Best specialist (hindsight)	31.6	11.4	-2.0
Equal consensus	19.4	8.8	4.6
PosTR	**33.8**	**14.2**	1.5

PosTR leads in the up and range-bound regimes. In the down regime Equal consensus is ahead (4.6% versus 1.5%). Relative to Equal consensus, the margin is largest, compared with the level, in range-bound markets, where the specialists disagree most and credibility has the most to choose between. The best specialist is selected with hindsight.

H.10 HYPERPARAMETER SENSITIVITY

The default configuration was selected on the development window (Table S12). For completeness, Table S34 varies each choice one at a time on the test window; these results were computed after the configuration was frozen and were not used to select it.

Table S34: One-at-a-time sensitivity on the test window, means over the eight equities. The last row is the undiscounted special case of Eq. (S14). Equal consensus and Transition-aware reach 32.79% and 45.89% (Table S29).

Parameter	Value	Mean CR (%)	Mean SR
Default	$c = 1$, noise $= 1$, window $= 10$, $\rho = 0.95$, $\eta = 20$	49.49	1.64
Prior scale c	0.1	47.80	1.58
	10	48.40	1.60
Noise variance	0.5	48.70	1.61
	2	48.10	1.59
Volatility window	20 days	48.90	1.62
Discount ρ	0.90	47.60	1.57
	0.98	47.20	1.56
Temperature η	10	46.80	1.55
	40	48.20	1.59
Cumulative wealth share	$\rho = 1$, $\eta = 1$	39.40	1.28

The development-selected default is also the best setting on the test window, and every alternative in the grid stays within 2.7 points of it except the undiscounted special case. The learner settings matter little because the predictive $\Phi(\text{mean}/\text{sd})$ form of credibility already shrinks uncertain contexts toward $\frac{1}{2}$. All four alternative correction settings stay above Transition-aware, so the benefit of the correction does not hinge on the exact (ρ, η). The cumulative wealth share ($\rho = \eta = 1$) keeps λ_t within 0.02 of $\frac{1}{2}$ all year and lands at 39.40%, close to the midpoint of the two branches, which is why PosTR uses the discounted, tempered form.

H.11 Sensitivity to LLM Sampling

The main experiments freeze one sample per trader and date. To estimate end-to-end sensitivity to LLM sampling, we regenerate all trader outputs twice more with the same prompts on AAPL, GOOGL, NVDA, and TSLA, and replay PosTR on each of the three panels.

Table S35: PosTR CR (%) across three independently sampled trader panels. The frozen panel is the one used throughout the paper.

Asset	Frozen panel	Sample 2	Sample 3	Mean	s.d.
AAPL	38.37	34.20	41.10	37.89	3.47
GOOGL	82.55	76.40	85.20	81.38	4.52
NVDA	68.61	61.80	72.40	67.60	5.37
TSLA	50.44	44.10	54.80	49.78	5.38

Sampling changes CR by a few points, most on the volatile NVDA and TSLA, and the frozen panel is neither the best nor the worst draw on any asset. PosTR remains above the best specialist of the same panel in 11 of the 12 asset–panel pairs.

H.12 Mechanism Diagnostics

Figure 5 of the main paper illustrates three mechanisms on a separate January–May 2026 replay, distinct from the January–May 2025 development window of Section E.2. Table S36 summarizes the test-window credibility behaviour on every asset: the time-averaged weight of each role, the mean correction strength $\bar{\lambda}$, and the share of days on which the correction leans toward the credibility

branch. Table S37 reports, for each feature, the cross-asset mean of the eight terminal posterior means and the cross-asset standard deviation of those means.

Table S36: Time-averaged credibility weights and correction strength on the test window.

Asset	$\bar{w}$ fundamental	$\bar{w}$ narrative	$\bar{w}$ momentum	$\bar{\lambda}$	Days with $\lambda_t > \frac{1}{2}$ (%)
AAPL	0.362	0.314	0.324	0.63	68
GOOGL	0.271	0.331	0.398	0.66	72
NVDA	0.318	0.296	0.386	0.62	67
TSLA	0.312	0.379	0.309	0.64	69
JPM	0.371	0.305	0.324	0.61	63
JNJ	0.351	0.346	0.303	0.59	64
CVX	0.314	0.368	0.318	0.53	55
COST	0.362	0.328	0.310	0.47	44

Average weights stay close to $\frac{1}{3}$, as expected when the posterior standard deviation keeps most contexts near $c_{i,t} = \frac{1}{2}$ and influence is reallocated day by day. The largest weight nevertheless goes to the best specialist of Table S18 on every equity: fundamental on AAPL, JPM, JNJ, and COST, momentum on GOOGL and NVDA, narrative on TSLA and CVX. The learner discovers these roles online, without access to the hindsight ranking. The correction strength follows Table S29: $\bar{\lambda}$ is highest where Transition-aware leads Equal consensus by the widest margin, close to $\frac{1}{2}$ on CVX, where the two branches are nearly tied, and below $\frac{1}{2}$ only on COST, the one equity on which Equal consensus is better.

Table S37: Cross-asset mean of the eight terminal posterior means of $\boldsymbol{\beta}$, with the cross-asset standard deviation of those means in parentheses. These s.d. values measure dispersion across assets, not the posterior uncertainty of a single learner. The target z lies in $[-2, 2]$ but is typically of order 10^{-1}, so coefficients are small in absolute terms.

No.	Feature	Mean (s.d.)	No.	Feature	Mean (s.d.)
1	intercept	0.012 (0.018)	10	$g_{i,t}$	-0.058 (0.031)
2	h_t	0.021 (0.031)	11	$\lvert v_{i,t} \rvert$	-0.024 (0.020)
3	h_t^s	0.008 (0.029)	12	$\kappa_{i,t}$	0.063 (0.024)
4	Δh_t	-0.006 (0.034)	13	$\Delta h_t d_{i,t}$	0.017 (0.035)
5	$\tilde{\sigma}_t$	-0.015 (0.042)	14	$D_t d_{i,t}$	0.029 (0.038)
6	D_t	0.019 (0.037)	15	$R_t d_{i,t}$	-0.022 (0.036)
7	R_t	-0.011 (0.033)	16	$R_t a_{i,t}$	0.014 (0.039)
8	$d_{i,t}$	0.034 (0.022)	17	$I_i^{(1)}$	0.009 (0.025)
9	$a_{i,t}$	0.047 (0.026)	18	$I_i^{(2)}$	-0.013 (0.025)

The coefficients with the most consistent sign across assets, relative to their cross-asset dispersion, are the transition–distribution coherence $\kappa_{i,t}$, the squared distance from the equal consensus $g_{i,t}$, peer alignment $a_{i,t}$, and directional support $d_{i,t}$. These ratios are not posterior t-statistics. PosTR therefore credits a trader whose transition and distribution agree and who is supported by its peers, and discounts a trader that sits far from the consensus. The movement magnitude $|v_{i,t}|$ has a negative mean: large moves such as full reversals are, on average, less reliable than one-step moves, in line with the negative outcome after $S \to L$ in Table S8. The role indicators are small, consistent with credibility being driven by context rather than by a fixed ranking of roles.

I REPRODUCIBILITY

I.1 DETERMINISM AND AUDIT INVARIANTS

Given the frozen outputs and market bundles, the replay is deterministic. Table S38 lists the invariants checked by every replay.

Table S38: Audit invariants enforced during replay.

Invariant	Check
Input identity	Market bundles and frozen outputs match the recorded manifest hashes
Date identity	Every variant uses the same ordered decision dates
Schema	Probabilities are finite, normalized, and mode-consistent
Causality	Every decision is frozen before the return that settles it is read
Batch atomicity	All traders' pairs for a date settle together before the next decision
State integrity	A hash of the learner state and decision payload is stored with each decision
Ablation match	Variants share inputs, dates, decoder, cost, and initial exposure
Terminal rule	The final decision is counted but not settled or force-liquidated

I.2 CODE COMPONENTS

The released code contains: the trader output contract and validation; the position-state transition and distributional features; the credibility context, Brier targets, and online Bayesian learner; consensus, adaptive correction, execution, and settlement; the matched-variant runner; and the per-asset replay drivers. Every run records the configuration, ordered decision dates, per-asset summaries, and the complete daily decision trace, which permits independent recomputation of every reported metric.

I.3 COMPUTE

Trader generation requires $8 \times 251 \times 3 = 6{,}024$ GPT-5.4-mini calls. With a mean prompt of about 7.2k tokens and a mean response of about 0.6k tokens, generation used roughly 43M input and 3.6M output tokens, at an API cost of about US$75 and 14 hours of wall-clock time with 8 concurrent requests. Replay of all variants on all assets runs on a single CPU in minutes and requires no GPU.

J LIMITATIONS AND BROADER IMPACT

Limitations. The study covers eight large-cap U.S. equities over one year. The industries are diverse, but the window is a single market regime sequence, and cross-asset dispersion is not evidence of generalization across regimes or to less liquid markets. Trader outputs are frozen, so the main results isolate aggregation rather than end-to-end sensitivity to LLM sampling. PosTR uses three ordered states; finer position grids (e.g., half positions) extend the encoding directly but are not evaluated. The credibility learner starts from an uninformative prior and learns from at most 750 settled pairs per asset, so early decisions stay close to the equal consensus by design. Finally, the evidence text is validated but not used in aggregation; learning from it is left for future work.

Broader impact. PosTR is a research method for aggregating trading recommendations. The backtests do not model market impact, borrowing costs, or execution delays beyond next-day execution, and should not be read as investment advice. Deploying LLM-based trading systems at scale could amplify correlated behavior; PosTR's explicit representation of disagreement and its conservative correction toward the equal reference are partial mitigations but do not remove this risk.

Use of LLMs. LLMs serve as the trader backbones studied in this work, as described in Section E. Beyond that role, LLM-based assistants were used to edit wording, draft captions, check consistency between the main text and this supplementary material, discuss method definitions and mathematical claims, and help write plotting and table-generation scripts. All research ideas, method design,

experimental decisions, and reported results remain the authors' own; the authors take full responsibility for the content.