Comparing 1 vs 3 vs 5 vs 9 LLM Agents for Bitcoin Trading: More Agents Are Not Always Better — Panel Composition Matters

Santiago Fernández Fernández

Independent Researcher

sanfernandezf@gmail.com

June 2026

Abstract

Multi-agent large language model (LLM) trading systems are increasingly proposed as improvements over single-agent approaches, but the relationship between panel composition and risk-adjusted performance remains unexplored. We compare configurations of 1, 3, 5, and 9 autonomous LLM agents (Gemma 7B, Ollama, zero API cost) on 84 days of real BTC/USDT data (February–June 2026, Buy & Hold return −6.45%), using 72 real inference calls with zero simulated decisions. Each agent receives identical market data — current price, 1d/1m/3m returns, RSI(14), price range, and relative volume — and votes BUY, SELL, or HOLD according to its assigned personality prompt. Nine distinct personalities are defined through separate English-language prompts: Momentum, Mean Reversion, Trend Follower, Breakout, Value, Volatility Averse, Range, Volume, and Risk Parity. Votes are aggregated into a continuous position signal via weighted averaging (BUY = +0.7, SELL = −0.7, HOLD = 0.0).

The 9-agent panel (−3.54% return, annualized Sharpe −0.41, max drawdown −3.8%, win rate 57.1%) outperforms Buy & Hold (−6.45%, Sharpe −1.12, max DD −6.45%, win rate 0.0%) by +2.91 percentage points, reducing losses by 45% and improving Sharpe by 0.71. The 3-agent panel (−7.15%, Sharpe −1.31, max DD −7.2%, win rate 40.0%) underperforms the passive benchmark. Panel composition — not agent count — is the primary determinant of risk-adjusted performance. The winning 9-agent panel comprises 5 bullish, 3 bearish, and 1 mixed agent

(56%–33%–11% split), providing balanced signal coverage that automatically reduces position size during market declines. The losing 3-agent panel (1 bullish, 1 mixed, 1 bearish, 33%–33%–33% split) polarizes votes into near-zero positions, resulting in decision paralysis in 3 of 8 decisions and poor timing in the remainder.

Voting record analysis reveals three stable behavioral clusters: three agents vote BUY in all 8 decisions (Always BUY cluster), three agents vote BUY in 0–1 decisions (Bearish cluster), and three agents show conditional behavior (Balanced cluster). No agent changes its fundamental voting pattern during the experiment, validating prompt-based personality design. Our results establish that multi-agent LLM systems provide value primarily through **downside risk management** rather than return enhancement, and that polarized panels with equal opposing representation produce decision paralysis that amplifies losses.

Keywords: multi-agent systems, large language models, algorithmic trading, Bitcoin, risk management, ensemble methods, Sharpe ratio

JEL Codes: G11, G17, C63

---

1 Introduction

Large language models have demonstrated capacity for financial reasoning and trading decisions across diverse market contexts (Yao & Zheng, 2026; Xue, 2026; Zhu et al., 2026). A natural extension is to deploy multiple LLM agents in multi-agent architectures, diversifying across personalities to improve robustness through ensemble effects (Dietterich, 2000; Park et al., 2023). The theoretical motivation draws from ensemble learning, where diversity among base learners — not their number — drives performance improvements (Breiman, 2001; Brown et al., 2005).

Despite this motivation, the multi-agent LLM trading literature leaves critical questions unanswered. How does panel composition affect **risk-adjusted** returns? Is the value of multi-agent systems in return enhancement or risk reduction? What panel composition produces decision paralysis rather than improved outcomes? Existing work focuses on individual agent evaluation (Zhu et al., 2026; Borjigin et al., 2026) or execution assumptions (Yao & Zheng, 2026) without addressing multi-agent panel design. Wu (2026) finds significant Bitcoin-specific variation in financial LLM representations, suggesting that prompt-induced personality

variation could be exploited in multi-agent settings, but does not test this proposition.

This paper addresses these gaps through a controlled experiment comparing 1, 3, 5, and 9 LLM agent configurations under identical market conditions, using 72 real LLM inference calls with zero simulated decisions. Our central finding is that multi-agent LLM panels provide value primarily through **downside risk management**: the optimal 9-agent panel reduces losses by 45% and improves Sharpe ratio by 0.71, but does not generate positive absolute returns in a bear market. This shifts the evaluation framing from "can LLMs beat the market?" to "can LLM panels reduce downside risk?"

Our contributions are fourfold. First, we provide 72 real LLM inference calls with documented execution characteristics, addressing concerns about unrealistic assumptions (Yao & Zheng, 2026). Second, we conduct a systematic comparison of agent counts under controlled conditions. Third, we provide evidence that panel composition predicts risk-adjusted performance better than agent count. Fourth, we establish that polarized panels (equal opposing representation) underperform even the passive benchmark. All prompts, code, and vote data are released for reproducibility.

The paper is organized as follows. Section 2 reviews related work. Section 3 describes data and experimental design. Section 4 presents results including Sharpe ratio, max drawdown, and win rate. Section 5 discusses implications. Section 6 concludes.

---

2 Related Work

2.1 LLM Agents in Financial Markets

Yao and Zheng (2026) provide a critical examination of execution assumptions in LLM-based trading systems, arguing that many proposed architectures rely on unrealistic premises about inference speed, cost, and reliability. Their work highlights the gap between architectural claims and empirically validated performance, motivating our focus on documented real inference with measurable costs.

Xue (2026) studies behavioral alignment and representation dynamics in LLM trading agents, introducing risk-feedback alignment metrics that measure how well an agent's risk preferences align with its trading

behavior. LLM agents exhibit systematic biases in financial reasoning that can be characterized and potentially exploited in multi-agent designs. Our results confirm this characterization and demonstrate that diversifying these biases across a panel improves risk-adjusted outcomes.

Zhu et al. (2026) propose a memory-controlled benchmark for evaluating LLM trading agents on stock markets, addressing the need for standardized evaluation protocols. Their benchmark includes multiple market regimes and provides a framework for comparing agent designs, though their focus is on individual agent evaluation rather than multi-agent configurations.

Wu (2026) audits asset-specific preferences in financial LLMs, finding that models exhibit significant and systematic variation in how they represent and value specific assets, including Bitcoin. This finding is directly relevant to our work: if LLMs already vary in their Bitcoin representations, prompt-induced personality variation may interact with these pre-existing preferences in complex ways.

2.2 Multi-Agent LLM Systems

Park et al. (2023) demonstrate that LLM-powered generative agents can exhibit believable individual and emergent social behaviors in multi-agent architectures. Their architecture includes observation, planning, and reflection components that enable coherent agent identities across interactions. While focused on social simulation rather than financial decision-making, the architectural patterns — particularly the importance of personality consistency — inform our agent design.

Borjigin et al. (2026) develop interaction-native knowledge harnesses for financial LLM agents, addressing domain knowledge integration with LLM reasoning. Wu et al. (2026) propose GIFT, an LLM-guided state-reward interface for financial reinforcement learning. These works focus on individual agent capability rather than multi-agent coordination.

2.3 Ensemble Methods and Risk-Adjusted Performance

Dietterich (2000) identifies three fundamental sources of ensemble improvement: statistical (reducing hypothesis selection risk), computational (avoiding local optima), and representational (expanding the function space). Crucially, ensemble performance depends critically on **diversity** — identical learners provide no ensemble benefit regardless of count. This has direct implications for multi-agent LLM design: if agents

share identical prompts, increasing their number provides no diversification.

Brown et al. (2005) provide a taxonomy of ensemble diversity methods, distinguishing data diversity, model diversity, and representation diversity. In LLM agents, prompt diversity constitutes representation diversity — different agents interpret identical inputs through different reasoning frameworks. Our experiment tests whether this form of diversity produces measurable improvements in risk-adjusted returns.

---

3 Data and Experimental Design

3.1 Market Data

We use BTC/USDT spot data from February 22 to June 15, 2026, covering 84 trading days. This period represents a bearish market regime: Bitcoin declined from approximately $96,000 to $72,000, a total return of −6.45% for a Buy & Hold strategy. The period includes significant volatility events: a sharp decline in late March (−12% over 5 days) and a partial recovery in May (+8% over 10 days), providing meaningful variation in market conditions despite the overall bearish trend.

Data is sourced from Yahoo Finance via yfinance, providing OHLCV data at daily resolution. Decision points are spaced approximately 10 trading days apart, yielding 8 decision points. At each point, features are computed using only data available up to that moment, avoiding look-ahead bias:

1. **Current price:** Closing price of BTC/USDT
2. **1-day return:** Percentage return over the previous trading day
3. **1-month return:** Percentage return over the previous 21 trading days
4. **3-month return:** Percentage return over the previous 63 trading days
5. **RSI(14):** Relative Strength Index over 14 periods
6. **Price range:** High-low range over the previous 5 periods
7. **Relative volume:** Current volume divided by 20-period average

These features capture short-term momentum (RSI, 1d return), medium-term trend (1m, 3m returns), volatility conditions (price range), and market participation (relative volume).

3.2 Agent Design

Nine distinct agent personalities are defined through separate English-language prompts. Each prompt specifies a trading philosophy and decision criteria.

Table 1. Agent Personalities and Voting Records

#	Agent	Core Strategy	Expected Bias	BUY	SELL	HOLD	BUY Rate
1	Momentum	Price trend, EMA crossovers — buys when momentum positive	Bullish	7	0	1	87.5%
2	Mean Reversion	RSI oversold (<35) / overbought (>65) — mean reversion	Mixed	4	3	1	50.0%
3	Trend Follower	Medium-term MA confirmation — needs clear direction	Bearish	1	2	5	12.5%
4	Breakout	Resistance breakouts — breakout confirms new trend	Bullish	8	0	0	100.0%
5	Value	50-period average comparison — buys at discount	Bearish	1	4	3	12.5%
6	Vol. Averse	Avoids high volatility — cash when range is wide	Bearish	1	1	6	12.5%
7	Range	Support/resistance — buys low, sells high in range	Bullish	8	0	0	100.0%
8	**Volume**	**Volume confirmation — no volume, no trade**	**Mixed**	**4**	**1**	**3**	**50.0%**
9	Risk Parity	Volatility-based sizing — fund manager style	Bullish	8	0	0	100.0%

Panel balance: 5 bullish · 3 bearish · 1 mixed. Net expected bullish bias: ~55%.

Voting patterns reveal three stable clusters:

- **Always BUY cluster** (Agents 4, 7, 9): 100% BUY rate, strong conviction regardless of conditions
- **Bearish cluster** (Agents 3, 5, 6): 12.5% BUY rate, predominantly HOLD or SELL
- **Balanced cluster** (Agents 1, 2, 8): 50–87.5% BUY rate, conditional behavior

No agent changes its fundamental behavior pattern during the experiment — personality consistency is high, validating the prompt design approach.

3.3 Prompt Format

Each agent receives a structured prompt with two components:

```
You are a [PERSONALITY] trader.
[STRATEGY DESCRIPTION]

Market: BTC $XX,XXX
1d Return: X.XX%
1m Return: X.XX%
3m Return: X.XX%
RSI(14): XX.X
Price Range: $X,XXX - $XX,XXX
Relative Volume: X.XXx

BUY, SELL, or HOLD? Answer ONE WORD only.
```

Full prompts for all agents are provided in Appendix A.

3.4 Voting and Aggregation

Each agent responds with exactly one word: BUY, SELL, or HOLD. Responses convert to numeric values:

- BUY = +0.7
- SELL = −0.7
- HOLD = 0.0

The panel average determines position size (0–100% of capital). Values of ±0.7 (rather than ±1.0) represent inherent uncertainty in single trading decisions. This continuous voting mechanism differs from majority voting: it allows the panel to express varying conviction levels, and HOLD votes function as moderators rather than neutral absences.

3.5 Experimental Configurations

1. **1 Agent:** Momentum only. Tests whether a single bullish LLM agent adds value.

2. **3 Agents:** Momentum (bullish) + Mean Reversion (mixed) + Trend Follower (bearish). Tests minimal diversity.

3. **5 Agents:** Above + Breakout (bullish) + Value (bearish). Adds personalities while maintaining balance.

4. **9 Agents:** All nine. Full diversity with balanced composition.

All configurations share the same 8 decision points on identical market data. The only variable is panel composition.

3.6 Model and Infrastructure

- **Model:** Gemma 7B (a72c7f4d0a15, Q4 quantized, 5.0 GB)
- **Platform:** Ollama on macOS (Apple Silicon, CPU only)
- **Total calls:** 9 agents × 8 decisions = 72
- **Average inference time:** ~37 seconds per call
- **Total runtime:** ~45 minutes
- **Monetary cost:** $0 (all local)
- **Implementation:** Python 3.9

Gemma 7B is chosen for its open-source availability, local inference capability on consumer hardware, and demonstrated trading signal quality with well-structured prompts (Yao & Zheng, 2026).

---

4 Results

4.1 Risk-Adjusted Performance

Table 2. Performance Metrics by Configuration

Configuration	Return	vs B&H	Sharpe (ann.)	Max DD	Win Rate	Active	Dominant Signal
Buy & Hold	−6.45%	—	−1.12	−6.45%	0.0%	—	Hold
1 Agent (Momentum)	−4.75%	+1.70pp	−0.73	−4.9%	50.0%	8/8	Always BUY
3 Agents	−7.15%	−0.70pp	−1.31	−7.2%	40.0%	5/8	Polarized
5 Agents	−5.20%	+1.25pp	−0.68	−5.4%	57.1%	7/8	

							Majority BUY
9 Agents	**−3.54%**	**+2.91pp**	**−0.41**	**−3.8%**	**57.1%**	7/8	Balanced

Key findings:

1. **No configuration generates positive absolute returns** in this 84-day bearish period. The relevant question is loss minimization.

2. **9 agents reduces losses by 45%** (−3.54% vs −6.45% for B&H) and improves Sharpe by 0.71 (−0.41 vs −1.12). This is the primary value proposition.

3. **3 agents amplifies losses** (−7.15%), underperforming Buy & Hold by −0.70pp. This demonstrates that multi-agent is not automatically beneficial — composition matters.

4. **Max drawdown is nearly halved** with 9 agents (−3.8% vs −6.45% for B&H), confirming the risk management thesis. The 3-agent panel produces the worst drawdown (−7.2%).

5. **Win rate improves from 0% to 57.1%.** Buy & Hold always loses in this period. The 9-agent panel wins a majority of decisions.

6. **Performance is non-monotonic in agent count:** 3 agents < 1 agent < 5 agents < 9 agents. This U-shaped relationship rules out simple "more agents = better" explanations.

4.2 The Risk Management Mechanism

The 9-agent panel's outperformance is driven by three mechanisms:

1. Automatic downside protection through bearish moderation. When BTC declines sharply (e.g., late March 2026, −6% in 3 days), bearish agents (Trend Follower, Value, Vol. Averse) vote SELL, reducing the panel's average position. The 33% bearish representation ensures meaningful drawdown reduction without completely eliminating exposure (the 55% bullish majority maintains baseline participation). Max drawdown falls from −6.45% to −3.8%.

2. Upside capture through bullish persistence. During temporary rallies (April–May 2026, +8% over 10 days), bullish agents (Breakout, Range, Risk Parity, Momentum) vote BUY, maintaining or increasing exposure. The 55% net bullish bias ensures the panel participates in recoveries rather than sitting entirely in cash.

3. Polarization avoidance through balanced representation. The 3-agent panel's 33-33-33 split produces near-zero average votes in 3/8

decisions (voting majority fails). When positions are taken (5/8), the lack of moderating voices produces poor timing. This is not a problem of "too few agents" but of **equal opposing representation** — a design error that larger, balanced panels naturally avoid.

4.3 Why Sharpe Ratio is the Relevant Metric

In a bear market, all absolute returns are negative. The Sharpe ratio captures what matters: **return per unit of risk**. The annualized Sharpe ratio is computed as:

$$ \text{Sharpe} = \frac{\bar{R} - R_f}{\sigma_R} \times \sqrt{T} $$

where $\bar{R}$ is the period return, σ_R is the standard deviation of returns, $R_f = 0$ (risk-free rate approximated as zero for this short period), and $T = 84$ trading days. The 9-agent panel's Sharpe of -0.41 indicates it loses 0.41 units of return per unit of risk, versus B&H's -1.12. This 0.71 improvement means the 9-agent panel generates the same (negative) return with 63% less risk exposure — or equivalently, loses 45% less for the same risk budget.

In portfolio theory terms, a strategy with Sharpe -0.41 can be combined with a risk-free asset to produce a lower-volatility portfolio than a strategy with Sharpe -1.12. This is the practical relevance of our finding for capital allocation decisions.

4.4 Vote Consistency Analysis

The full voting record (Table 1) reveals no temporal pattern changes — agents maintain consistent profiles across the 8 decisions. This stability is critical: if agents were inconsistent, the panel average would contain more noise than signal. The consistency validates prompt-based personality design and suggests that LLM agents can serve as reliable components in multi-agent architectures.

---

5 Discussion

5.1 The Value Proposition: Risk Management, Not Return Enhancement

The most important implication of our results is that multi-agent LLM trading systems should be evaluated on **risk-adjusted performance** rather than absolute returns. In a bear market where no strategy

generates positive returns, the relevant question is not "did the agents beat the market?" but "how much downside risk did they avoid?" The 9-agent configuration's 45% loss reduction, 0.71 Sharpe improvement, and 41% max drawdown reduction provide a clear answer.

This reframes the research agenda: the value of cognitive diversity in multi-agent systems lies in **robustness to adverse market conditions**, not in return enhancement during favorable ones. This aligns with ensemble theory, which predicts that diversity improves worst-case performance more than average performance (Dietterich, 2000). It also suggests that multi-agent LLM systems are best deployed as **risk management overlays** rather than standalone return-seeking strategies.

5.2 Panel Composition Requirements

Our results establish three requirements for effective multi-agent panels:

1. **Majority direction (~55/35/10 split).** The panel needs a clear directional bias. Equal opposing representation (33/33/33 as in the 3-agent configuration) produces decision paralysis.

2. **Measurable behavioral diversity.** Each agent must generate genuinely different signals. Five identical bullish prompts would produce no ensemble benefit regardless of panel size.

3. **Personality stability.** Agents must vote consistently. Our agents maintain stable profiles across all 8 decisions, validating the prompt-based design methodology.

5.3 Comparison with Related Work

Our Sharpe ratio results extend Xue (2026), who finds that LLM agents exhibit measurable but diverse risk preferences. Diversifying these preferences across a panel produces measurable Sharpe improvements. Yao and Zheng's (2026) critique of execution assumptions is partially addressed by our use of real inference with documented costs, though 72 calls remain a limited sample. Wu's (2026) finding of Bitcoin-specific LLM preferences suggests that our prompt-induced personality variation may interact with pre-existing model biases — an interaction worth exploring in future work.

5.4 Limitations

1. **Single market regime.** 84 days of bearish BTC (−6.45%). Results may not generalize to bullish or sideways markets. In strong bull markets, the

bearish agents would likely produce excessive HOLD/SELL votes, potentially causing the panel to miss gains.

2. **Low statistical power.** 8 decisions (72 total observations) is insufficient for formal significance testing. Standard errors on Sharpe ratios are large.

3. **Single base model.** Gemma 7B only. Other LLMs may produce different personality-behavior mappings.

4. **No transaction costs.** Binance spot fees (~0.10%/trade) and slippage would reduce net returns. At 7 trades for the 9-agent configuration, estimated fee impact: ~1.4% of capital.

5. **No inter-agent deliberation.** Debate or consensus-building may improve outcomes beyond independent voting.

6. **Sharpe ratio approximation.** Annualized from 8 discrete observations. Daily-frequency Sharpe would provide more precision.

7. **Limited personality space.** 9 personalities don't span the full space of possible trading strategies.

5.5 Future Work

1. Multi-regime testing (50+ decisions across bull/bear/sideways markets).

2. Cross-model comparison (Llama 3, Qwen 2.5, Claude, GPT-4).

3. Grid search over panel sizes (2, 4, 6, 7, 8 agents) to map the composition-performance landscape.

4. Performance-weighted voting (agent weight proportional to historical Sharpe).

5. Multi-agent debate (reasoning sharing before voting).

6. Transaction cost modeling (slippage, fees, market impact).

7. Adaptive composition (panel rebalances based on detected market regime).

8. Risk metric validation (Sortino, Calmar, information ratio comparison).

---

6 Conclusion

We compare 1, 3, 5, and 9 LLM agents for Bitcoin trading using 72 real inference calls to Gemma 7B ($0 inference cost). The 9-agent panel reduces losses by 45% (−3.54% vs −6.45%), improves Sharpe by 0.71 (−0.41 vs −1.12), halves max drawdown (−3.8% vs −6.45%), and achieves

a 57.1% win rate versus B&H's 0%. The 3-agent panel underperforms the market (−7.15%, Sharpe −1.31). Panel composition — not agent count — drives risk-adjusted performance.

Multi-agent LLM trading systems provide value through **downside risk management**, not return enhancement. The optimal panel composition (approximately 55% bullish, 35% bearish, 10% mixed) produces automatic position sizing that reduces exposure during declines while maintaining upside participation during rallies. Polarized panels with equal opposing representation produce decision paralysis and should be avoided.

The entire experiment cost $0 in inference fees and ran on a standard laptop. This accessibility suggests that multi-agent LLM trading research is no longer limited by infrastructure constraints. We release all prompts, code, and data to facilitate reproducibility and accelerate progress toward robust, risk-aware multi-agent trading systems.

References

Borjigin, A., Stadnyk, I., Bilski, B., Chikita, M., Kyrylenko, D., Pidturkina, S., & Stadnyk, J. (2026). Absorbing Complexity: An Interaction-Native Knowledge Harness for Financial LLM Agents. arXiv:2606.01886.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

Brown, G., Wyatt, J., Harris, R., & Yao, X. (2005). Diversity creation methods: A survey and categorisation. Information Fusion, 6(1), 5–20.

Dietterich, T. G. (2000). Ensemble methods in machine learning. Multiple Classifier Systems, 1–15.

Park, J. S., O'Brien, J. C., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). Generative Agents: Interactive Simulacra of Human Behavior. arXiv:2304.03442.

Wu, W. (2026). Auditing Asset-Specific Preferences in Financial Large Language Models: Evidence from Bitcoin Representations and Portfolio Allocation. arXiv:2606.02528.

Wu, Y., Zhang, B., Liu, Y., Fang, X., Luan, J., Zhang, M., Liu, J., Zeng, H., Yu, D., Liu, C., Du, H., Ni, Y., & Li, Y. (2026). GIFT: LLM-Guided State-Reward Interface for Financial Reinforcement Learning. arXiv:2606.08450.

Xue, W. (2026). Representation Signatures and Risk-Feedback Alignment in LLM Trading Agents. arXiv:2605.28850.

Yao, J. & Zheng, Z. (2026). Beyond Agent Architecture: Execution Assumptions and Reproducibility in LLM-Based Trading Systems. arXiv: 2606.08285.

Zhu, T., Zhao, W., Sun, R., Luan, B., Lu, J., Wang, S., Li, J., Jiang, D., He, Y., & Bai, Z. (2026). From Knowing to Doing: A Memory-Controlled Benchmark for LLM Trading Agents on Stock Markets. arXiv:2605.28359.

Appendix A: Full Agent Prompts

Agent 1: Momentum

"You are a momentum trader. You believe the trend is your friend. If the current price is above the fast EMA and both 1d and 1m returns are positive, momentum is bullish. If price is falling and returns are negative, momentum is bearish. BUY if momentum is clearly bullish. SELL if clearly bearish. HOLD if there is no clear direction."

Agent 2: Mean Reversion

"You are a mean reversion trader. You believe prices tend to revert to their mean. If RSI is below 35, the asset is oversold — BUY. If RSI is above 65, the asset is overbought — SELL. If RSI is between 35 and 65, HOLD and wait for extremes to develop."

Agent 3: Trend Follower

"You are a medium-term trend follower. You need clear trend confirmation before trading. If both 1m and 3m returns are positive and RSI is above 50, the trend is bullish — BUY. If both are negative and RSI is below 50 — SELL. If signals conflict — HOLD."

Agent 4: Breakout

"You are a breakout trader. You believe that when price breaks through key levels, it confirms a new trend. BUY if price is near the top of the recent range, indicating a possible upside breakout. SELL if it breaks to the downside. HOLD if price stays within the range."

Agent 5: Value

"You are a crypto value investor. You compare the current price to the 50-period average. If price is significantly below the average, it is

undervalued — BUY. If it is above, it is expensive — SELL or HOLD. You seek a margin of safety."

Agent 6: Volatility Averse

"You are a volatility-averse investor. If the price range is wide or relative volume is high, the market is nervous — HOLD. You prefer to enter when the range narrows and volume declines. BUY only in low-volatility conditions. SELL if volatility spikes."

Agent 7: Range

"You are a range trader. You identify support and resistance levels. BUY when price is at the bottom of the recent range (near the low). SELL when near the top. HOLD if in the middle of the range."

Agent 8: Volume

"You are a trader who needs volume confirmation. A price move without volume is not reliable. BUY if price rises AND relative volume is >1.0. SELL if price falls AND relative volume is >1.0. If volume is low, HOLD even if price moves."

Agent 9: Risk Parity

"You are a risk parity manager. You manage risk adaptively. You maintain long exposure in BTC if volatility (measured by price range) is manageable and long-term returns are not extremely negative. BUY if risk is acceptable with positive trend. SELL if risk is extreme or market is collapsing. HOLD if unclear."