arXiv:2602.00948v1 [physics.soc-ph] 1 Feb 2026

FINEVO：从孤立回测到多智能体金融策略演化的生态市场博弈

邹明希[1], [2*] **陈家祥**[1], [*], [‡] **罗傲天**[1] **戴静怡**[1] **张驰**[2]**孙东宁**[3], [†]**徐增林**[1], [†] [1]复旦大学[2]Tensorpacific[3]鹏城实验室{mxzou24, jiaxiangchen23}@m.fudan.edu.cn

摘要

传统的金融策略评估依赖于静态环境中的孤立回测。这类评估独立考察每项策略，忽略相关性与相互作用，也无法解释为何某些策略最终会在演化市场中存续或消亡。我们转向生态视角，将交易策略建模为在共享市场中相互作用并学习的自适应智能体。我们并非提出一种新策略，而是提出**FinEvo**，一种用于研究多智能体金融策略演化动力学的生态博弈形式化框架。在个体层面，不同类型的基于机器学习的交易者——基于规则、深度学习、强化学习以及大语言模型（LLM）智能体——利用历史价格和外部新闻等信号进行适应。在群体层面，策略分布通过三种专门设计的机制——选择、创新和环境扰动——不断演化，以捕捉真实市场中的动态力量。这两个适应层面相结合，将演化博弈论与现代学习动力学联系起来，为研究策略行为提供了有原则的环境。针对外部冲击和真实世界新闻流开展的实验表明，FinEvo 既稳定、可复现，又能灵活揭示依赖语境的结果。策略可能占据主导、走向崩溃或形成联盟，具体取决于其竞争者——这些模式在静态回测中无从显现。通过将策略评估重新定义为生态博弈形式化框架，FinEvo 提供了一套统一的机制层面方案，用于分析多智能体金融市场中的稳健性、适应和涌现动力学，并有望为探索宏观经济政策与金融监管对价格演化和均衡的潜在影响提供途径。

1 引言

金融市场是最复杂、最具适应性的系统之一，其结果不仅取决于外部信号，也取决于参与者的集体行为。大多数评估范式都将这类互动抽象掉了。深度学习强调预测准确率，从基于注意力机制的架构（如 TimeMixer（Wang 等人，2024年））到基础模型（如 Chronos（Ansari 等人，2024年））皆是如此。强化学习拓展至序贯决策，近期进展包括离线强化学习，以实现数据高效的策略学习（Ku- mar 等人，2020年）；以及风险敏感强化学习，以提升不确定性下的稳健性（Duan 等人，2016年）。与此同时，基于大语言模型的智能体通过推理和工具使用拓展交易逻辑：FinCon 将交易构建为结合 CVaR 的结构化金融推理（Yu 等人，2024年）；FinGPT 调整基础模型以用于预测和决策支持（Liu 等人，2023年）；TradingAgents 则探索通用规划和成交能力（Xiao 等人，2024年）。

[*]共同贡献。
[†]通讯作者。
[‡]项目负责人。

然而，大多数现有方法都在孤立环境中进行优化，忽略了塑造现实世界市场表现的生态相互依赖关系。我们的目标并非提出新的交易策略或预测模型，而是建立一套严谨的生态博弈形式体系——将其作为“金融风洞”，在受控环境中对系统性风险和市场冲击进行压力测试（见附录 H）。

ABIDES（Byrd 等，2020年）和 Mesa（Masad 等，2015年）等仿真基础设施支持市场微观结构研究和基于智能体的实验，但它们主要作为成交环境运行，并未对种群层面的适应或长期战略动态进行建模。PyMarketSim（Mascioli 等，2024年）进一步为深度强化学习智能体提供高保真的订单簿执行，但仍以单智能体优化为中心，而非关注集体生态行为。TraderTalk（Vidler 和 Walsh，2024年）探索由大语言模型驱动的行为交易智能体，但侧重双边互动，缺乏用于市场系统性评估的框架。

除这些平台之外，JAX-LOB（Frey 等，2023年）等 GPU 加速的限价订单簿模拟器，以及面向交易延迟和强化学习的快速基于智能体的框架（Belcak 等，2021年），推动了基于历史订单流数据的高保真训练；与此同时，HFTE（Mahdavi-Damghani，2017年）等生态系统式模型则研究高频策略之间的多物种互动。然而，这些研究方向均未提供**统一的进化博弈论公式化表述**，以刻画异质交易生态中的适应、种群流动和系统性影响——FinEvo 正是针对这一空白展开研究（见表 6）。

为实现这一愿景，我们提出**FinEvo**，这是一种用于研究多智能体金融策略演化动态的生态博弈形式体系。智能体群体涵盖基于规则、深度学习、强化学习和大语言模型（LLM）交易者；每个交易者都在参数化策略和资本约束下运行，价格则由连续双向拍卖清算内生决定。受复制子—突变方程（Nowak，2006年）的启发，FinEvo通过选择、创新和环境扰动对种群动态进行形式化建模，刻画适应、灭绝以及优势循环和状态转变等生态系统层面的现象，从而能够在真实的市场互动中，以严谨的方式评估稳健性和适应能力。我们在附录 A 中对相关工作进行了更详细的综述。

总而言之，我们的贡献如下：

1）我们提出**FinEvo**，这是一种将异质交易者与演化种群动态相结合的生态博弈形式体系，并通过**FinEvo 随机微分方程**对其进行形式化描述。其机制分解将选择、创新和环境扰动统一起来，使评估从单个实例的收益提升至机制层面的 行为。2）在温和条件下，我们为 FinEvo 随机微分方程确立了最低限度的理论保证——单纯形不变性、正性以及存在性与唯一性，并推导出宏观层面方差分解，将系统波动性归因于选择、创新和扰动。3）通过针对冲击和现实世界信号开展实验，我们证明 FinEvo 能够揭示静态回测无法捕捉的依赖情境的表现、稳健性和涌现的联盟动态。

这些贡献共同确立了 FinEvo 作为研究多智能体金融市场中稳健性与适应的严谨基础。

2 生态博弈：问题表述与演化动力学

我们将金融策略评估建模为一种生态博弈，其中 N 异质交易代理群体在一个共享且动态演化的市场中相互作用。每个智能体遵循独特的策略，并通过市场结果与外部信号之间的反馈进行适应。这一表述使我们能够同时研究个体策略的表现，以及选择、创新和扰动的集体动力学。

代理行为与市场反馈。 在每个时间步 t，智能体 i 选择动作 $a_{i,\,t} \sim \pi_{k(i)}(\mathcal{I}_t)$，其中 $\pi_{k(i)}$ 是类型为 k的策略，以近期价格 $\{P_{t-\tau}, \ldots, P_t\}$ 和外生信号 E_t为条件。动作空间包括市价买入、市价卖出、限价买入、限价卖出和持有。

市场将所有动作和冲击 ξ_t 汇总为下一时刻的价格 P_{t+1}，详细的撮合机制见附录 G.6，

$$P_{t+1} = F\big(P_t, \{a_{i,t}\}_{i=1}^{N}, \xi_t\big).$$

随后，已执行交易会更新投资组合持仓。具体而言，令$C_{i,t}$ 分别表示智能体 i 在时间 t的现金持仓，$S_{i,t}$ 表示其资产持仓，而 P_t 表示市场出清价格。$C_{i,t+1} = C_{i,t} - a_i^{\text{buy}}{}_{,t}\, P_t + a_i^{\text{sell}}{}_{,t}\, P_t$，以及$S_{i,t+1} = S_{i,t} + a_i^{\text{buy}}{}_{,t} - a_i^{\text{sell}}{}_{,t} \cdot$ 对于策略 π_k，其在$[t, t+1]$期间的实际收益为$\hat{f}_{k,t} = \frac{1}{N_k}$，$\sum_{i:\pi_{k(i)}=\pi_k} \Big[C_{i,t+1} + S_{i,t+1}P_{t+1} - (C_{i,t} + S_{i,t}P_t)\Big]$，代表采用者的平均增量财富变化。

个体适应。每种策略都维护一个内部参数状态$\Psi_{k,t}$，该状态决定其如何将信息 $\mathcal{I}_t$ 映射为动作。此状态通过一个通用的适应算子不断演变

$$\Psi_{k,t+1} = \mathcal{A}\big(\Psi_{k,t}, \hat{f}_{k,t}, E_t\big),$$

其中，$\mathcal{A}$概括了多种学习和调整机制。在实践中，这种更新可能对应于强化学习中的时序差分或策略梯度规则、深度学习模型中的基于梯度的参数调优、大型语言模型中的上下文内适应，或基于规则的策略中的启发式调整。

除了对即时奖励做出反应，交易员通常还会尝试估计其策略的长期价值。为此，我们为每个 π_k 关联一个前瞻价值过程 $V_k(t)$，其模型为 $dV_k(t) = \lambda_k\big(f_k(E_t) - V_k(t)\big)\,dt + \nu_k\,dB_k(t)$，其中 $f_k(E_t)$ 表示受环境条件影响的目标收益，$\lambda_k > 0$ 控制调整速度，而 $\nu_k dB_k(t)$ 引入随机探索。这个奥恩斯坦–乌伦贝克过程描述了长期价值估计向外部信号收敛，同时保留由不确定性带来的变化性，并作为前瞻性锚点指导后续的种群层面动力学。

群体演化（FinEvo 随机微分方程）。设 $X_t = (x_1(t) \ldots, x_K(t))^\top \in \Delta^{K-1}$ 表示在 K 种策略上的种群分布，其中 $\sum_k x_k(t) = 1$。X_t 的演化可以抽象地表示为一个算子

$$x_{t+1} = \mathcal{G}\big(x_t, V(t), m_t, \xi_t\big),$$

其中，$V(t) = (V_1(t) \ldots, V_K(t))^\top$ 汇集了在个体层面估计的前瞻性价值，m_t 表示创新分布，ξ_t 表示随机扰动。

具体而言，我们将 $\mathcal{G}$ 表述为一个连续时间随机微分方程，即**FinEvo 随机微分方程**，它将三种设计机制——选择、创新和扰动——整合为统一动态：

$$\begin{aligned} dX_t &= b(X_t, t)\,dt + \Sigma(X_t, t)\,dW_t \\ &= \underbrace{\beta \mathrm{diag}(X_t)\,\big(V(t) - \bar{V}(t)\mathbf{1}\big)}_{\textbf{Selection}}\,dt + \underbrace{\mu\,(m_t - X_t)}_{\textbf{Innovation}}\,dt + \underbrace{\gamma\,\mathrm{diag}(X_t)\,\sigma\,P(X_t)\,dW_t}_{\textbf{Perturbation}}, \end{aligned}$$

其中，$\bar{V}(t) = X_t^\top V(t)$为群体平均收益，$P(X_t) = I - \mathbf{1}X_t^\top$将扰动投影到单纯形切空间上，$\sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_K)$，$\gamma$表示全局扰动缩放因子，$W_t$是一个 K维布朗运动。目标分布 $m^t \sim \mathrm{Dir}(\alpha_1, \ldots, \alpha_K)$对社会影响和随机试验进行建模，从而确保 $m^t \in \Delta^{K-1}$并维持不同策略之间的多样性。

这种形式保留了单纯形，确保正性，并在温和条件下存在不变测度（附录 E）。虽然与复制子—突变方程相关，但 FinEvo 将选择建立在收益信号 $V(t)$ 的基础上，并明确纳入创新和扰动，从而构建出更灵活、更稳健的群体演化模型。

独立性假设。我们在方差分解中采用独立性假设（A1–A3），仅仅是为了得到闭式表达式。对于 FinEvo 随机微分方程本身，这些假设并非 必需：即使收益与环境冲击相关，动力学也不会改变。在相关情形下，分解只需增加协方差项（见附录 D）。

机制分解。 在离散近似中，份额 Δx_k 的短期变化由三种力量塑造：
$\Delta x_k \approx \beta x_k^t\big(V_k^t - \bar{V}^t\big) + \mu\big(m_k^t - x_k^t\big) + x_k^t\Big(\gamma\sigma_k\eta_k^t - \gamma\sum_j x_j^t\sigma_j\eta_j^t\Big), \qquad \eta_k^t \sim \mathcal{N}(0,\ 1)$。

相应地，它们对短期波动率的贡献为

$$\text{方差}(\Delta x_k) \approx \underbrace{\beta^2 (x_k^t)^2\,\frac{\nu_k^2}{2\lambda_k}}_{\text{选择}} + \underbrace{\mu^2\,\frac{m_k(1-m_k)}{\alpha_0+1}}_{\text{创新}} + \underbrace{\gamma^2\Big((x_k^t)^2\sigma_k^2 - 2x_k^t\sigma_k\sum_j x_j^t\sigma_j + \Big(\sum_j x_j^t\sigma_j\Big)^2\Big)}_{\text{扰动}}.$$

这些分解共同展示了选择、创新和扰动如何共同塑造金融市场中进化适应的方向和波动率。详细证明见附录 C。

3 演化动力学的实验与分析

3.1 实验设置

为研究金融策略如何在生态环境中演化，我们构建了一个多智能体仿真环境。在该环境中，我们测试了两种情景：一种包含人工冲击，另一种包含现实世界新闻事件。我们仿真了一个包含新闻冲击、多指标再分布和生态进化更新的合成市场（见附录 G.6.1）。关于这些情景动机的完整讨论见附录 G.7。

智能体群体 市场由 20 种交易智能体原型组成，每种原型都有不同的行为、信息来源和决策逻辑。为便于说明，我们将这些智能体分为四大类：（1）**基于规则的智能体**，包括趋势跟踪、均值回归（Jegadeesh 和 Titman，1993年；De Bondt 和 Thaler，1985年）和噪声交易等传统策略，以及基本面（信息型）交易者；我们还为其赋予不同的交易风格（例如，激进型和保守型），以体现现实世界的多样性。（2）**深度学习智能体**，即 Informer 和 TimeMixer 等监督模型*Informer* 和 *TimeMixer*（Zhou 等人，2021年；Wang 等人，2024年），它们使用神经网络，根据市场指标预测价格方向。（3）**强化学习（RL）智能体**，例如 DoubleDQN*DoubleDQN*和近端策略优化*PPO*（Van Hasselt 等人，2016年；Schulman 等人，2017年）等自适应策略，它们通过与市场交互进行在线学习，以最大化累积收益。（4）**大语言模型（LLM）智能体**，它们解析新闻和公告等文本信息，以推断市场情绪并生成交易信号（例如，*FinCon*、*TradingAgents*）（Yu 等人，2024年；Xiao 等人，2024年），并以 GPT-4o-mini（Hurst 等人，2024年）作为底层语言模型。各智能体设计的详细信息见附录 F.2.1。正文中通篇采用缩写，完整对照表汇总于附录 F.1，完整的实验参数配置见附录 G.6.2。

模拟场景 每种配置均通过蒙特卡洛实验进行评估，每种配置进行**128 次独立运行** 。我们考虑两种设置：（1）**人工冲击**，向市场注入预定义的扰动，以考察系统稳定性和韧性（见附录 G.6.4）；（2）**真实世界新闻**，使用实际历史新闻流（例如 GDELT 和 Reuters Markets；详情见附录 G.6.5）模拟2023年7月至2025年7月。市场每天在09:00至17:00进行连续双向拍卖（时长略长于实际交易时段，以便市场状态趋于稳定），使事件驱动智能体和大语言模型智能体能够处理真实信息，并在符合现实的日内约束下进行交互。

分析指标 我们的评估框架采用微观–中观–宏观结构（详见附录 G.6.3）。在**微观层面**，我们通过群体占比（各智能体的市场份额）、智能体胜率（智能体的收益在评估试验中排名前三的频率），以及包括带置信区间的收益、夏普比率、最大回撤和换手率在内的金融指标，评估个体表现与多样性。在**中观层面**，我们通过集中度（赫芬达尔-赫希曼指数）、策略熵、模块度（社区强度）、协同作用与对抗（相关性热图）、

共现频率，以及反映依赖关系和联盟重新配置的互信息网络，来刻画种群多样性与交互结构。在**宏观层面**，我们通过将波动率分解为绩效压力（V_{select}）、创新／社会混合（$V_{\text{innovation}}$）和环境扰动（$V_{\text{perturbation}}$），并测量相变（用于捕捉主导策略簇的变化），评估系统波动性和状态转变。正式定义见附录 G.6.3。为便于可视化，我们使用 Savitzky–Golay 滤波器对原始时间序列数据（如群体占比和波动性指标）进行平滑处理。此外，遵循实证金融领域的标准做法，我们报告超额峰度、偏度、夏普比率离散度和波动率的波动率，以量化肥尾、收益不对称性、跨策略表现离散度和波动率聚集。这些指标广泛用于刻画真实市场的典型事实，可为 FinEvo 生成的内生动力学的真实性提供外部验证。

3.2 人工冲击下的仿真

我们测试市场生态在大型外部人工扰动下的适应情况。在我们的合成市场中，我们在 12:00 至 13:30 之间注入新闻冲击，同时考虑强正面和强负面两种情景。该设置使我们能够考察微观、中观和宏观层面的适应。

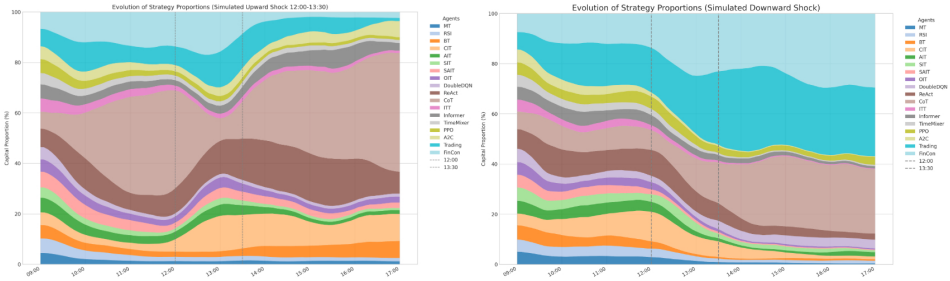
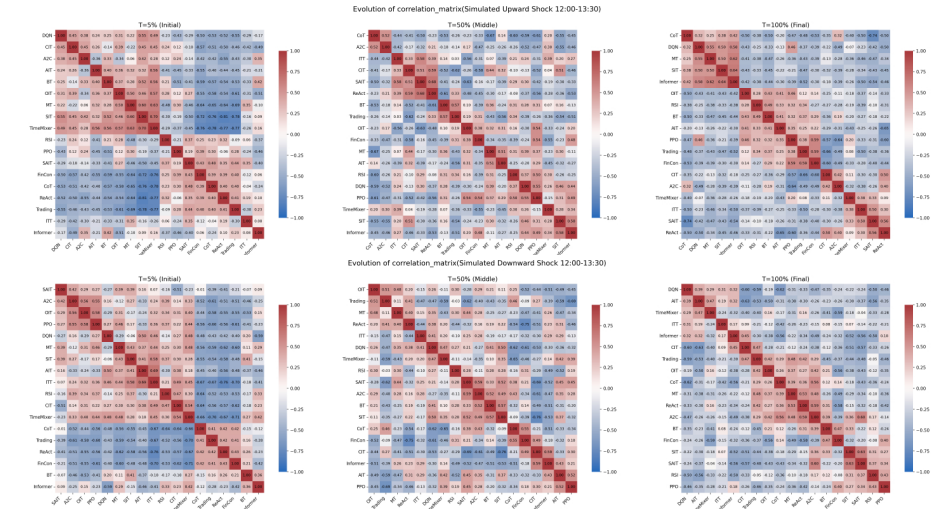


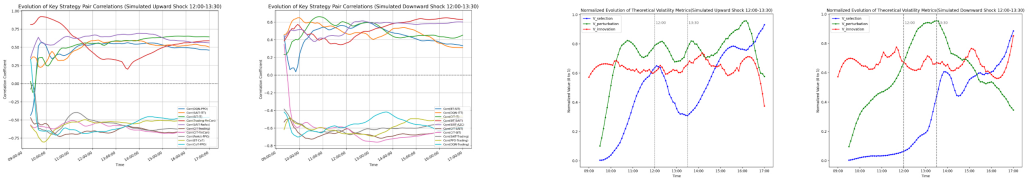
图 1：智能体群体占比随时间的变化。左：正向冲击；右：负向冲击。

微观：市场份额动态。 图 1 显示，市场的响应存在显著的不对称性。正向冲击通过削弱思维链的主导性，并将市场份额重新分配给多个智能体（例如 CIT、ReAct），提高了多样性；而负向冲击则加剧集中度，形成由大语言模型主导的均衡（FinCon、ReAct）。因此，冲击会驱动主导地位集中度发生制度依赖型转变，揭示了静态基准无法捕捉的动态变化。



Figu图 2：正向冲击（上）和负向冲击（下）下相关矩阵随时间的变化。红色 竞争，蓝色 =。

中观：交互结构与稳定性。 图 2 展示了联盟的重组过程。冲击发生前，大型红色簇表明存在许多策略协同行动的协调群体。冲击期间，这些簇逐渐分化，蓝色带状区域开始显现，反映出更激烈的竞争和更大的分歧。冲击过后，新的合作簇形成，但其结构各不相同：正向情形下更趋多元，负向情形下则集中于大语言模型周围。这一对比凸显了适应具有路径依赖性。



(a) 冲击下的策略对相关性。　　(b) 冲击下的波动性分解。

图3：外部冲击下的响应：左侧 = 正向冲击，右侧 = 负向冲击。在 (b) 中，蓝色表示 = 选择，红色表示 = 创新，绿色表示 =扰动。

深入观察各个策略对（图3a），我们发现了三种一致的模式：(1) 一些策略对保持持续合作，(2) 另一些策略对保持持续竞争，以及 (3) 冲击仅引发局部重组。因此，尽管联盟会发生变化，更广泛的合作—竞争骨干结构仍保持完整，从而提供结构稳定性。

宏观：波动性分解。 图3b将波动率分解为三个组成部分：选择（$\mathbf{V}_{\text{selection}}$）、创新（$\mathbf{V}_{\text{innovation}}$）和扰动（$\mathbf{V}_{\text{perturbation}}$）。在两种情景下，冲击都会触发扰动激增，使联盟趋于不稳定。随后，选择占据主导地位：在正向情景下，它支撑多元化均衡；而在负向情景下，它推动系统趋向大语言模型主导地位。创新保持稳定，作为一种背景力量维持多样性，而非驱动短期变化。

摘要。 外部冲击会在各个层面重塑市场：它们会打破垄断或加剧集中度（微观），瓦解并重组联盟（中观），并在选择恢复秩序之前放大波动率（宏观）。总体而言，适应受选择、创新与扰动之间的相互作用支配。冲击会在短期内加剧扰动，但长期均衡主要由选择塑造，创新则维持多样性。

为求完整，我们还分析了冲击下共现矩阵和互信息网络的演化，这从网络交互的角度为稳健性提供了补充证据。详细结果见附录G.1。

3.3 日内演化动力学的实证分析

在考察模拟冲击下的适应性动力学之后，我们接下来转向真实世界数据，分析单个交易日内的**日内演化动力学**。这使我们能够验证，在受控实验中观察到的生态演化模式是否会在实际市场条件下持续存在。

日内适应与盈利能力。图4表明，日内动态遵循多阶段周期：强选择最初将资本集中于主导策略，持续的创新随后恢复多样性与平衡，临近收盘时加剧的扰动打破联盟，随后新的均衡出现。Informer与TimeMixer最终胜出，凸显进化压力如何在单个交易日内共同塑造状态转变。除生存之外，一个自然的问题是：哪些智能体能够在市场摩擦下将适应转化为盈利能力。表1报告了相同轮次的平均值（收益及95%置信区间、夏普比率、最大回撤、换手率）。大多数智能体出现亏损，但少数智能体尽管回撤较大仍实现正收益（以阴影标示），这得益于保持多样性的创新。更高的换手率进一步表明，盈利主体交易更为活跃，凸显出一种既具选择性又具包容性的市场生态。

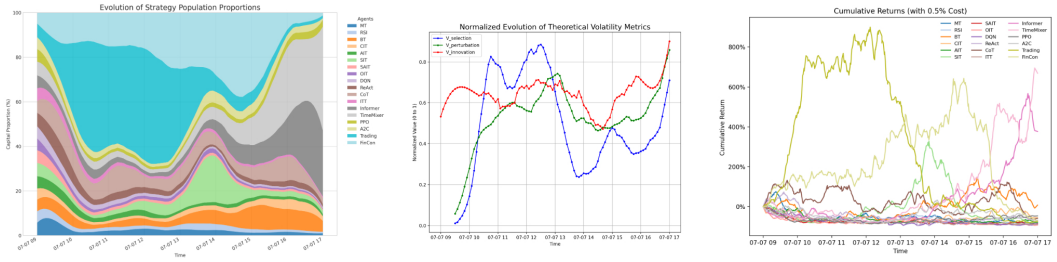
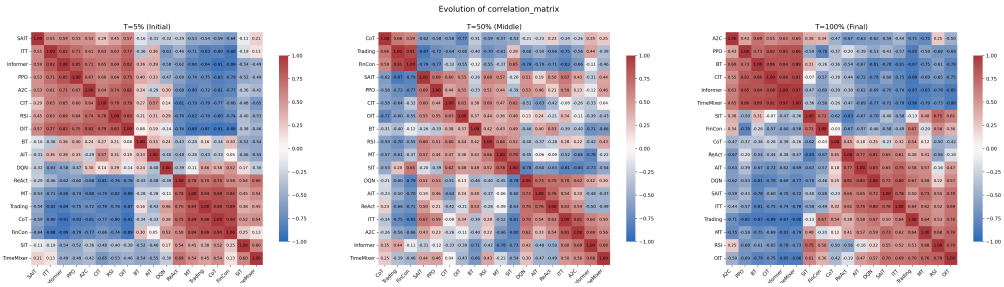


图4：2025年7月7日的日内演化：（左）智能体群体比例，（中）理论波动性指标，（右）在0.5%交易成本（含滑点）下的累计收益。

表1：按智能体划分的绩效指标（收益及95%置信区间、夏普比率、最大回撤、换手率；2025年-07-07，N=128）。正收益以灰色阴影标示。

智能体	收益	夏普比率	最大回撤	换手	智能体	收益	夏普比率	最大回撤	换手	智能体	收益	夏普比率	最大回撤	换手
MT	-0.832 ± 0.064	-2.016	-0.908	0.157	ReAct	-0.685 ± 0.025	15.972	-0.978	0.207	In	4.333 ± 0.022	26.928		-0.771	0.362
RSI	-0.849 ± 0.023	-3.546	-0.888	0.122	思维链	-0.560 ± 0.046	5.118	-0.815	0.434	Ti	6.738 ± 0.031	30.392		-0.790	0.463
BT	0.034 ± 0.010	10.804	-0.802	0.309	ITT	-0.896 ± 0.066	-1.228	-0.910	0.119	近端策略优化	-0.431 ± 0.033	6.641		-0.859	0.123
CIT	-0.544 ± 0.032	6.265	-0.844	0.163	交易	-0.746 ± 0.048	-2.244	-0.979	0.890	优势演员-评论家	-0.733 ± 0.004	0.847		-0.837	0.197
AIT	-0.807 ± 0.026	-3.121	-0.838	0.126	FinCon	-0.685 ± 0.045	-2.605	-0.958	0.845	深度 Q 网络	-0.885 ± 0.034	13.436		-0.955	0.139
SIT	-0.673 ± 0.057	0.624	-0.948	0.334	SAIT	-0.866 ± 0.010	-0.933	-0.909	0.099	OIT	-0.915 ± 0.032	2.265		-0.942	0.130

动态联盟重构。除了个体结果外，我们接下来还将考察交互结构层。图5展示了智能体间相关性的日内演化，揭示了协同联盟 与战略对抗之间的动态交互。市场起初呈现多个竞争集群，且内部协调性较强；随后，在环境扰动加剧的作用下，这些集群逐渐分裂并重新组织。临近收盘时，这些短暂的变化最终汇聚成一种新的市场状态，其特征是联盟重构以及力量平衡的改变。我们还在附录G.1中报告了共现矩阵和互信息网络的日内演化。



Fig图5：2025年7月7日日内相关矩阵的动态演化。红色块表示 t *syn*协同联盟，而蓝色色块表示战略对抗。

牛市与熊市中的表现排名。最后，从日内波动转向市场周期，我们评估这些模式是否会在长期牛市和熊市状态中持续存在。我们将2022年1月至10月近似为一个熊市阶段，其客观界定依据是合成环境中持续的下行动态，并将2024年10月至2025年7月近似为一个牛市阶段，其特征是持续复苏并不断创出新高。图6比较了这些市场状态中的领奖台名次与前三名胜率。

我们的分析表明，智能体表现高度**依赖市场状态**：牛市会使少数策略占据主导地位，而熊市则会形成竞争格局更为分散的市场。值得注意的是，**基于大语言模型的智能体在两种市场状态下的排名始终更高**，这表明，它们整合多样化信息线索并灵活调整的能力，相较于基于固定规则或纯粹量化信号的策略，带来了显著的稳健性优势。

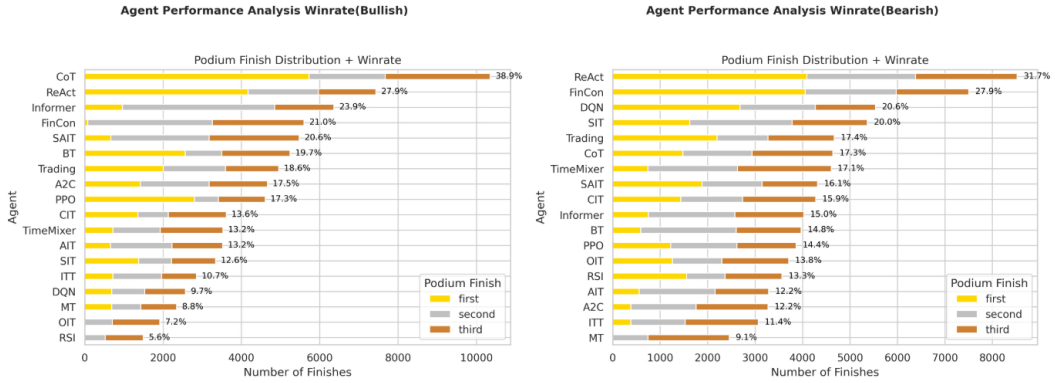


图6：牛市与熊市中的智能体表现排名。

表2：演化博弈建模与基于智能体的仿真中的生态指标比较

平台与阶段	熵$_{\text{mean}}$	[illegible]	前1$_{\text{mean}}$	平均绝对相关性$_{\text{mean}}$	模块度$_{\text{mean}}$	相变$_{\text{mean}}$
FinEvo（牛市）	0.732 ± 0.016	0.226 ± 0.032	0.453 ± 0.024	0.441 ± 0.019	0.091 ± 0.028	3 ± 0.437
FinEvo（熊市）	1.167 ± 0.036	0.175 ± 0.029	0.241 ± 0.067	0.376 ± 0.039	0.063 ± 0.018	4 ± 0.348
ABIDES（牛市）	0.516 ± 0.045	0.539 ± 0.011	0.746 ± 0.031	0.653 ± 0.046	0.153 ± 0.023	1 ± 0.439
ABIDES（熊市）	0.623 ± 0.011	0.481 ± 0.012	0.615 ± 0.045	0.541 ± 0.022	0.132 ± 0.032	1 ± 0.344
PyMarketSim（牛市）	0.539 ± 0.005	0.564 ± 0.019	0.773 ± 0.046	0.662 ± 0.039	0.159 ± 0.034	1 ± 0.332
PyMarketSim（熊市）	0.662 ± 0.037	0.461 ± 0.015	0.719 ± 0.023	0.575 ± 0.007	0.144 ± 0.019	1 ± 0.329

市场环境中的演化博弈建模与基于智能体的仿真 表2比较了 FinEvo、ABIDES 和 PyMarketSim 在牛市和熊市期间的表现。FinEvo 展现出一种动态、适应性强的环境，智能体会持续调整策略，从而促进更高的多样性和竞争。相比之下，ABIDES 和 PyMarketSim 更为僵化，随着时间推移，占据主导地位的策略较少，限制了适应性和多样性。这表明 FinEvo 在捕捉市场演化方面具有优势，而其他平台则表现出更为静态的行为。

3.4 消融研究

进化压力的消融分析 为区分三种进化压力（选择、创新、扰动）的作用，我们依次移除每个组成部分，并比较由此产生的种群动态（图7）。移除$\mathbf{V}_{\text{selection}}$（左图）后，竞争性强化消失，所有智能体的占比都保持相近，形成静态且低效的生态。移除$\mathbf{V}_{\text{innovation}}$（中图）后，多样性崩溃，因为 FinCon 和 Trading 迅速占据主导地位，导致寡头垄断锁定。移除$\mathbf{V}_{\text{perturbation}}$（右图）后，动态变化变得更平缓但也更僵化：FinCon 和 Trading 仍占主导地位，而自适应重组受到抑制。总体而言，选择推动效率，创新维持多样性，而来自环境的扰动可能会放大系统内的冲击，从而防止过早锁定并促成状态转变。

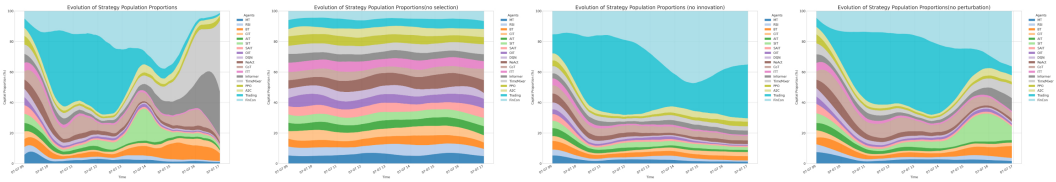


图7：进化压力的消融研究。从左到右：移除选择、创新和扰动。

我们通过全评估期间的汇总统计（表3）补充图7中的单日动态。这些全局指标提供了更稳健的图景：移除**选择** 会导致熵虚高、相变频繁，表明出现了不稳定碎片化；移除创

表3：全局生态指标的消融研究结果（均值 ± 95% 置信区间）。

设置	熵$_{\text{mean}}$	赫芬达尔-赫希曼指数$_{\text{mean}}$	前1$_{\text{mean}}$	平均绝对相关性$_{\text{mean}}$	模块度$_{\text{mean}}$	相变$_{\text{mean}}$
FinEvo	0.828 ± 0.14	0.189 ± 0.012	0.344 ± 0.015	0.384 ± 0.021	0.074 ± 0.009	4 ± 0.663
FinEvo（无选择）	1.793 ± 0.25	0.081 ± 0.108	0.074 ± 0.009	0.139 ± 0.018	0.121 ± 0.012	14 ± 0.767
FinEvo（无创新）	0.535 ± 0.13	0.477 ± 0.014	0.531 ± 0.033	0.553 ± 0.024	0.023 ± 0.007	1 ± 0.483
FinEvo（无扰动）	0.719 ± 0.04	0.245 ± 0.011	0.376 ± 0.017	0.474 ± 0.019	0.036 ± 0.008	3 ± 0.591

新 会使多样性崩溃，并将生态锁定在寡头垄断状态；移除**扰动** 会提高集中度，但减少状态转变，反映出系统的僵化。综合来看，这些结果证实了先前的案例研究：选择促进效率，创新保持多样性，而扰动则推动自适应重组。

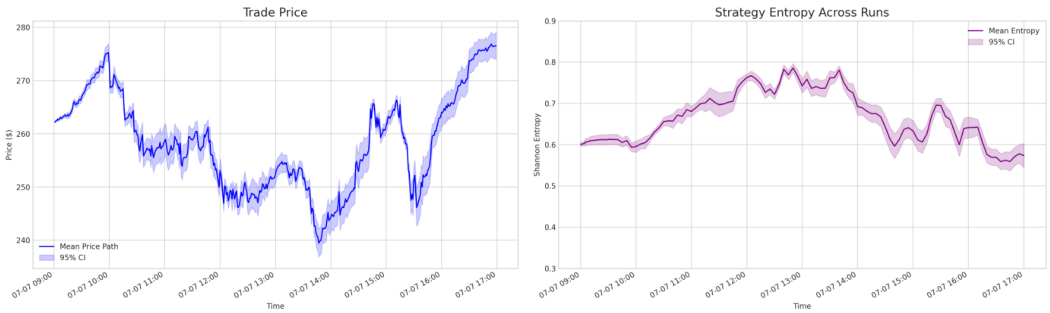


图8：对随机种子的稳健性（2025年7月7日）。左：交易价格轨迹；右：策略熵。曲线表示均值；阴影带表示基于128次运行计算的95%置信区间。

稳健性消融（随机种子）。 我们基于2025年7月7日进行的模拟开展了消融研究，模拟共进行128轮，旨在评估我们的框架在不同随机种子下的稳健性。如图8所示，我们报告了交易价格轨迹和策略熵的跨运行均值，并以阴影带表示各次运行的95%置信区间。

尽管各次运行之间存在轻微差异，核心模式仍保持一致：（i）从价格方面看，市场在外部冲击后呈现稳定的上升趋势，同时伴有有界波动；（ii）从策略方面看，熵持续处于较高水平，表明多样性和异质性得以维持。阴影区域相对狭窄，说明不同随机种子会产生相似的动态变化，从而证明我们的框架在随机扰动下具有稳定性和稳健性。此外，我们还报告了模拟外生冲击下的交易价格轨迹，进一步展现了框架的稳健性（详细结果见附录G.1）。

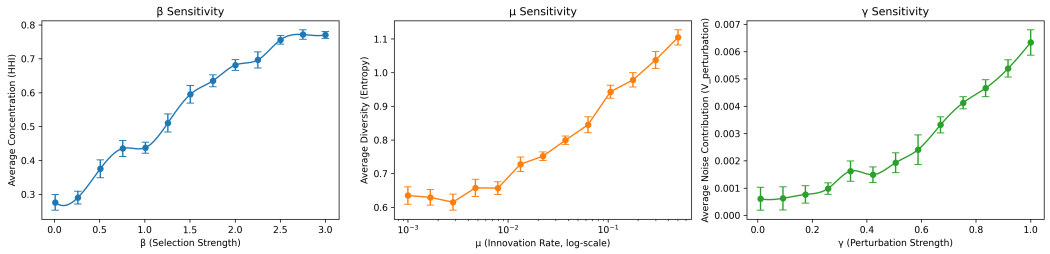


图9：演化参数的敏感性分析（2025年7月7日的仿真）。左：选择强度 β 对市场集中度的影响。中：创新率 μ 对策略多样性的影响。右：扰动强度 γ 对波动率贡献的影响。误差条表示128次运行的95%置信区间。

参数敏感性消融实验。 我们进一步针对进化框架的三个关键参数开展敏感性实验：选择强度 β、突变率 μ 和环境扰动强度 。

我们使用三个互补指标报告系统级结果：(i) 平均集中度（赫芬达尔–赫希曼指数，HHI）、(ii) 平均多样性（香农熵）和 (iii) 平均扰动贡献 $V_{\text{perturbation}}$。图9展示了结果。随着 β 增大，平均集中度显著上升，这表明更强的选择压力会加速少数策略占据主导地位。相反，较高的 μ 会提升多样性，证实突变和社会混合机制有助于维持异质性。同时，增大 γ 会使 $V_{\text{perturbation}}$ 单调上升，说明外生扰动对波动率具有独立贡献。此外，我们还进行了 $\pm 20\%$ 扰动参数稳健性测试；结果（附录 G.3）表明，所有典型事实指标均保持稳定，证实 FinEvo 的动态不受初始智能体参数影响。

市场规模消融实验 N。 我们考察 FinEvo 对市场规模 N 的稳健性，即参与该生态的智能体数量。表 4 报告了四种市场规模 $N \in \{60, 80, 100, 120\}$ 下的典型事实指标和绩效指标，涵盖牛市和熊市状态（每种设置运行 128 次）。

表 4：市场规模消融实验 N（128 次运行）。在不同 N 下，典型事实指标保持稳定。数值为均值 $\pm$ 95% 置信区间。

市场状态	N	超额峰度$_{\text{mean}}\uparrow$	偏度$_{\text{mean}}\downarrow$	夏普比率离散度$_{\text{mean}}\uparrow$	波动率的波动率$_{\text{mean}}\uparrow$
看涨	120	3.143 ± 0.141	-1.539 ± 0.132	1.673 ± 0.056	0.009 ± 0.0004
	100	3.163 ± 0.138	-1.532 ± 0.129	1.669 ± 0.061	0.008 ± 0.0004
	80	3.178 ± 0.165	-1.541 ± 0.158	1.657 ± 0.073	0.008 ± 0.0005
	60	3.159 ± 0.178	-1.516 ± 0.162	1.659 ± 0.067	0.007 ± 0.0005
看跌	120	4.326 ± 0.169	-1.805 ± 0.152	1.759 ± 0.091	0.012 ± 0.0006
	100	4.259 ± 0.172	-1.793 ± 0.149	1.732 ± 0.089	0.013 ± 0.0007
	80	4.115 ± 0.175	-1.765 ± 0.161	1.741 ± 0.138	0.013 ± 0.0007
	60	4.251 ± 0.197	-1.772 ± 0.183	1.725 ± 0.119	0.011 ± 0.0007

对于 N 的所有取值，FinEvo 始终保持厚尾损益（超额峰度较高）、显著的下行不对称性（负偏度），以及相近的夏普比率离散度和波动率的波动率；平均收益仅在狭窄区间内波动。

3.5 程式化事实验证

一个关键问题是，尽管存在高斯外生冲击，FinEvo 的内生价格形成能否再现经典市场程式化事实（Cont，2001）。为此，我们比较 **FinEvo**、**ABIDES** 和 **PyMarketSim** 的典型事实指标，在牛市和熊市状态下分别运行 128 次模拟，并计算标准计量经济学指标（超额峰度、偏度、夏普比率离散度和波动率的波动率）。与基线方法相比，FinEvo 更能匹配厚尾、负偏度和较高的波动率的波动（表 5），且这一表现具有稳健性：采用相关的厚尾冲击变体后，所得典型事实指标相似（附录 G.4，表 9）。我们还在附录 G.5 中进行了小规模的跨资产稳健性检验。

表 5：不同模拟器的典型事实指标比较（128 次运行）。数值报告均值 $\pm$95%置信区间。

模拟器 / 市场状态	超额峰度$_{\text{mean}}\uparrow$	偏度$_{\text{mean}}\downarrow$	夏普比率离散度$_{\text{mean}}\uparrow$	波动率的波动率$_{\text{mean}}\uparrow$
FinEvo（牛市）	3.163 ± 0.138	-1.532 ± 0.129	1.669 ± 0.061	0.008 ± 0.0004
FinEvo（熊市）	4.259 ± 0.172	-1.793 ± 0.149	1.732 ± 0.089	0.013 ± 0.0007
ABIDES（牛市）	0.315 ± 0.041	-0.209 ± 0.036	0.675 ± 0.028	0.004 ± 0.0002
ABIDES（熊市）	0.562 ± 0.053	-0.474 ± 0.047	0.836 ± 0.033	0.007 ± 0.0003
PyMarketSim（牛市）	0.446 ± 0.044	-0.186 ± 0.032	0.512 ± 0.026	0.003 ± 0.0002
PyMarketSim（熊市）	0.591 ± 0.058	-0.438 ± 0.041	0.906 ± 0.034	0.005 ± 0.0003

4 结论与未来工作

在本研究中，我们提出了一种用于市场生态建模的生态博弈形式体系，将选择、创新和环境扰动统一到单一数学表述中。我们在多智能体金融市场环境中实现了该框架，并通过大规模

模拟展示了该框架捕捉不同市场阶段中稳定性、集中度和多样性演变的能力，从而有助于缩小理论分析与经验动态之间的差距。总体而言，我们的研究建立了从机制到指标再到行为的三层映射，凸显了所提框架的可解释性和稳健性。FinEvo 采用模块化组件设计，可自然地支持进一步扩展。附录 I 简要讨论了未来可能的研究方向。

5 伦理声明

FinEvo 提供了一种生态博弈形式体系，用于研究金融策略的演化。所用数据集包括公开可用的金融市场数据和合成新闻流，不涉及个人或敏感信息，因此隐私方面的顾虑极少。为降低潜在风险，我们的框架旨在用于市场动态研究和策略评估，而非开发可部署的交易算法。

6 可复现性声明

我们致力于确保研究结果的可复现性。本研究使用的所有数据集均可公开获取（见附录 G.6.5），我们将在论文发表后公开源代码、模型实现和配置文件。为保证稳健性和稳定性，每项实验均进行了 128 次蒙特卡洛运行，并报告平均结果及 95%置信区间（见附录 G.6.1）。正文和附录中提供了模型架构 F.2.1、超参数 G.6.2 和评估指标 G.6.3 的详细说明，以便独立验证。

参考文献

Selim Amrouni、Aymeric Moulin、Jared Vann、Svitlana Vyetrenko、Tucker Balch 和 Manuela Veloso。ABIDES-Gym：用于多智能体离散事件仿真的 Gym 环境及其在金融市场中的应用。载于 第二届 *ACM* 金融人工智能国际会议论文集，第 1–9 页，2021年。

Torben G Andersen、Tim Bollerslev、Francis X Diebold 和 Paul Labys。已实现波动率的建模与预测。*Econometrica*，71(2)：579–625，2003年。

Abdul Fatir Ansari、Lorenzo Stella、Caner Turkmen、Xiyuan Zhang、Pedro Mercado、Huibin Shen、OleksandrShchur、SyamaSundar Rangapuram、Sebastian PinedaArango、Shubham Kapoor 等。Chronos：学习时间序列的语言。*arXiv* 预印本 *arXiv:2403.07815*，2024年。

Peter Auer、Nicolo Cesa-Bianchi 和 Paul Fischer。多臂老虎机问题的有限时间分析。*Machine Learning*，47(2)：235–256，2002年。

Thomas B¨ack、David B Fogel 和 Zbigniew Michalewicz。《进化计算手册》。发行版，97(1)：B1，1997年。

Nicholas Barberis、Andrei Shleifer 和 Robert Vishny。投资者情绪模型。金融经济学杂志，49(3)：307–343，1998年。

Peter Belcak、Jan-Peter Calliess 和 Stefan Zohren。应用于强化学习及交易延迟效应研究的快速基于智能体的仿真框架。载于国际多智能体系统与基于智能体仿真研讨会，第42–56页。施普林格，2021年。

Richard Bellman。动态规划。科学，153(3731)：34–37，1966年。

Tim Bollerslev。广义自回归条件异方差。计量经济学杂志，31(3):307–327，1986年。

Jean-Philippe Bouchaud、J Doyne Farmer 和 Fabrizio Lillo。《市场如何缓慢消化供需变化》。载于金融市场手册：动态与演化，第 57–160 页。爱思唯尔，2009年。

William Brock、Josef Lakonishok 和 Blake LeBaron。技术分析的盈利能力：美国股票市场中移动平均和交易区间突破规则的研究。《金融学期刊》，47(5)：1731–1764，1992年。S´ebastien Bubeck、Nicolo Cesa-Bianchi 等。随机与非随机多臂老虎机问题的遗憾分析。《*Machine Learning*》基础与趋势，5(1)：1–122，2012年。David Byrd、Maria Hybinette 和 Tucker Hybinette Balch。ABIDES：迈向高保真多智能体市场仿真。载于*2020*年 *ACM SIGSIM* 高级离散仿真原理会议论文集，第 11–22 页，2020年。Lili Chen、Kevin Lu、Aravind Rajeswaran、Kimin Lee、Aditya Grover、Misha Laskin、Pieter Abbeel、Aravind Srinivas 和 Igor Mordatch。决策变换器：通过序列建模实现强化学习。《神经信息处理系统进展》，34：15084–15097，2021年。Paul F Christiano、Jan Leike、Tom Brown、Miljan Martic、Shane Legg 和 Dario A modei。基于人类偏好的深度强化学习。《神经信息处理系统进展》，30，2017年。Rama Cont。资产收益的实证特性：典型事实与统计问题。《定量金融》，1(2)：223，2001年。Werner FM De Bondt 和 Richard Thaler。股票市场是否反应过度？《金融学期刊》，40(3)：793–805，1985年。Kalyanmoy Deb、Amrit Pratap、Sameer Agarwal 和 TAMT Meyarivan。一种快速且精英化的多目标遗传算法：NSGA-II。*IEEE* 进化计算汇刊，6(2)：182–197，2002年。Yan Duan、Xi Chen、Rein Houthooft、John Schulman 和 Pieter Abbeel。连续控制深度强化学习的基准测试。载于国际机器学习大会，第 1329–1338 页。PMLR，2016年。Robert F Engle。利用英国通货膨胀方差估计值研究自回归条件异方差。载于*ARCH*：精选读物。牛津大学出版社，1995年11月。ISBN 9780198774310。doi：10.1093/oso/9780198774310.003.0001。URL https://doi.org/10. 1093/oso/9780198774310.003.0001。Robert F Engle 和 Victor K Ng。衡量并检验新闻对波动率的影响。《金融学期刊》，48(5)：1749–1778，1993年。J Doyne Farmer 和 Duncan Foley。经济学需要基于智能体的建模。《自然》，460(7256)：685–686，2009年。Sascha Yves Frey、Kang Li、Peer Nagy、Silvia Sapora、Christopher Lu、Stefan Zohren、Jakob Foer-ster 和 Anisoara Calinescu。JAX-LOB：GPU 加速的限价订单簿模拟器，助力大规模强化学习交易。载于第四届 *ACM* 金融人工智能国际会议论文集，第 583–591 页，2023年。Bingzheng Gan、Yufan Z hao、Tianyi Zhang、Jing Huang、Yusu Li、Shu Xian Teo、Changwang Zhang 和 Wei Shi。Master：一个专精于大语言模型和蒙特卡洛树搜索的多智能体系统。*arXiv* 预印本 *arXiv:2501.14304*，2025年。Azul Garza、Cristian Challu 和 Max Mergenthaler-Canseco。TimeGPT-1。*arXiv* 预印本 *arXiv:2310.03589*，2023年。Evan Gatev、William N Goetzmann 和 K Geert Rouwenhorst。配对交易：相对价值套利规则的表现。《金融研究评论》，19(3)：797–827，2006年。Jean-Bastien Grill、Florian Strub、Florent Altch´e、Corentin Tallec、Pierre Richemond、Elena Buchatskaya、Carl Doersch、Bernardo Avila Pires、Zhaohan Guo、Mohammad Gheshlaghi Azar 等。自举潜在表征：一种自监督学习新方法。《神经信息处理系统进展》，33：21271–21284，2020年。

Campbell R Harvey、Edward Hoyle、Russell Korgaonkar、Sandy Rattray、Matthew Sargaison与Otto Van Hemert。波动率目标策略的影响。可于*SSRN 3175538*获取，2018年。Josef Hofbauer与Karl Sigmund。演化博弈与种群动态。剑桥大学出版社，1998年。Cars H Hommes。经济学与金融中的异质主体模型。《计算经济学手册》，2:1109–1186，2006年。Allen H Huang、Hui Wang与Yi Yang。FinBERT：用于从金融文本中提取信息的大语言模型。《当代会计研究》，40(2):806–841，2023年。Aaron Hurst、Adam Lerer、Adam P Goucher、Adam Perelman、Aditya Ramesh、Aidan Clark、AJ Ostrow、Akila Welihinda、Alan Hayes、Alec Radford等。GPT-4o系统卡。*arXiv*预印本 *arXiv:2410.21276*，2024年。Max Jaderberg、Valentin Dalibard、Simon Osindero、Wojciech M Czarnecki、Jeff Donahue、Ali Razavi、Oriol Vinyals、Tim Green、Iain Dunning、Karen Simonyan等。基于种群的神经网络训练。*arXiv*预印本 *arXiv:1711.09846*，2017年。Narasimhan Jegadeesh与Sheridan Titman。买入赢家并卖出输家的收益：对股票市场效率的启示。《金融学期刊》，48(1):65–91，1993年。Aviral Kumar、Aurick Zhou、George Tucker与Sergey Levine。面向离线强化学习的保守Q学习。《神经信息处理系统进展》，33:1179–1191，2020年。Albert S Kyle。连续拍卖与内幕交易。*Econometrica*：计量经济学会期刊，第1315–1335页，1985年。Blake LeBaron。基于主体的计算金融。《计算经济学手册》，2:1187–1233，2006年。Joel Z Leibo、Edward Hughes、Marc Lanctot与Thore Graepel。自适应课程以及创新从社会互动中的涌现：多智能体智能研究宣言。*arXiv*预印本 *arXiv:1903.00742*，2019年。Sergey Levine。作为概率推断的强化学习与控制：教程与综述。*arXiv*预印本 *arXiv:1805.00909*，2018年。Xiao-Yang Liu、Guoxuan Wang、Hongyang Yang与Daochen Zha。FinGPT：为金融大语言模型普及互联网规模数据。*arXiv*预印本 *arXiv:2307.10485*，2023年。Junyu Luo、Zhizhuo Kou、Liming Yang、Xiao Luo、Jinsheng Huang、Zhiping Xiao、Jingshu Peng、Chengzhong Liu、Jiaming Ji、Xuanzhe Liu等。FinMME：金融多模态推理评估基准数据集。*arXiv*预印本 *arXiv:2505.24714*，2025年。Thomas Lux与Michele Marchesi。金融市场随机多智能体模型中的尺度缩放与临界性。《自然》，397(6719):498–500，1999年。Babak Mahdavi-Damghani。介绍HFTE 模型：面向高频量化金融策略的多物种捕食者—猎物生态系统。《*Wilmott*》，2017(89):52–69，2017年。David Masad、Jacqueline L Kazil等。Mesa：基于主体的建模框架。载于*SciPy*，第51–58页，2015年。Chris Mascioli、Anri Gu、Yongzhao Wang、Mithun Chakraborty与Michael Wellman。使用深度强化学习的交易智能体金融市场模拟环境。载于第五届*ACM*金融人工智能国际会议论文集，第117–125页，2024年。Oliver Mihatsch与Ralph Neuneier。风险敏感型强化学习。*Machine Learning*，49(2):267–290，2002年。Volodymyr Mnih、Koray Kavukcuoglu、David Silver、Andrei A Rusu、Joel Veness、Marc G Bellemare、Alex Graves、Martin Riedmiller、Andreas K Fidjeland、Georg Ostrovski等。通过深度强化学习实现人类水平控制。《自然》，518(7540):529–533，2015年。

Volodymyr Mnih、Adria Puigdomenech Badia、Mehdi Mirza、Alex Graves、Timothy Lillicrap、Tim Harley、David Silver与Koray Kavukcuoglu。面向深度强化学习的异步方法。载于国际机器学习大会，第1928–1937页。PmLR，2016年。Martin A Nowak。演化动力学：探索生命的方程。哈佛大学出版社，2006年。Boris N Oreshkin、Dmitri Carpov、Nicolas Chapados与Yoshua Bengio。N-BEATS：面向可解释时间序列预测的神经基扩展分析。*arXiv*预印本 *arXiv:1905.10437*，2019年。Martin L Puterman。马尔可夫决策过程：离散随机动态规划。约翰·威立父子出版公司，2014年。William H Sandholm。群体博弈与演化动力学。麻省理工学院出版社，2010年。John Schulman、Filip Wolski、Prafulla Dhariwal、Alec Radford与Oleg Klimov。近端策略优化算法。*arXiv*预印本 *arXiv:1707.06347*，2017年。Richard S Sutton、Andrew G Barto等。强化学习：导论，第1卷。麻省理工学院出版社（剑桥），1998年。Peter D Taylor与Leo B Jonker。演化稳定策略与博弈动力学。数学生物科学，40(1-2):145–156，1978年。Karl Tuyls、Pieter Jan'T Hoen与Bram Vanschoenwinkel。重复博弈中多智能体学习的演化动力学分析。《自主智能体与多智能体系统》，12(1):115–153，2006年。Masatoshi Uehara、Jiawei Huang与Nan Jiang。面向离策略评估的极小极大权重与Q函数学习。载于国际机器学习大会，第9659–9668页。PMLR，2020年。Hado Van Hasselt、Arthur Guez与David Silver。基于双Q学习的深度强化学习。载于*AAAI*人工智能大会论文集，第30卷，2016年。Alicia Vidler与Toby Walsh。TraderTalk：应用于模拟人类双边交易互动的大语言模型行为主体仿真模型。载于*2024*年*IEEE*国际智能体大会（*ICA*），第164–167页。IEEE，2024年。Oriol Vinyals、Igor Babuschkin、Wojciech M Czarnecki、Micha¨el Mathieu、Andrew Dudzik、Junyoung Chung、David H Choi、Richard Powell、Timo Ewalds、Petko Georgiev等。利用多智能体强化学习在星际争霸II中达到宗师水平。《自然》，575(7782):350–354，2019年。Shiyu Wang、Haixu Wu、Xiaoming Shi、Tengge Hu、Huakun Luo、Lintao Ma、James Y Zhang与Jun Zhou。TimeMixer：面向时间序列预测的可分解多尺度混合。*arXiv*预印本 *arXiv:2405.14616*，2024年。Yijia Xiao、Edward Sun、Di Luo与Wei Wang。TradingAgents：多智能体大语言模型金融交易框架。*arXiv*预印本 *arXiv:2412.20138*，2024年。Yangyang Yu、Zhiyuan Yao、Haohang Li、Zhiyang Deng、Yuechen Jiang、Yupeng Cao、Zhi Chen、Jordan Suchow、Zhenyu Cui、Rong Liu等。FinCon：具有概念化语言强化、用于增强金融决策的合成式大语言模型多智能体系统。《神经信息处理系统进展》，37:137010–137045，2024年。Yangyang Yu、Haohang Li、Zhi Chen、Yuechen Jiang、Yang Li、Jordan W Suchow、Denghui Zhang与Khaldoun Khashanah。FinMem：具有分层记忆与角色设计的性能增强型大语言模型交易智能体。《*IEEE*大数据汇刊》，2025年。Haoyi Zhou、Shanghang Zhang、Jieqi Peng、Shuai Zhang、Jianxin Li、Hui Xiong与Wancai Zhang。Informer：超越高效Transformer的长序列时间序列预测。载于*AAAI*人工智能大会论文集，第35卷，第11106–11115页，2021年。Tian Zhou、Ziqing Ma、Qingsong Wen、Xue Wang、Liang Sun与Rong Jin。FEDformer：面向长期序列预测的频率增强分解式Transformer。载于国际机器学习大会，第27268–27286页。PMLR，2022年。

A 相关工作

A.1 金融智能体

金融智能体这一概念将市场参与者抽象为具有不同规则、目标和信息集的决策实体。这一框架强调：（i）异质性——智能体在趋势跟踪、均值回归或新闻敏感度等行为假设方面各不相同；（ii）交互——结果不仅取决于外部数据，也取决于其他参与者的动作。通过将策略视为智能体，既有研究提供了一种统一的抽象，可涵盖多种建模方法。

在这一视角下，已经出现了若干类别。基于规则的智能体包括动量、均值回归、突破和配对交易规则（Brock 等人，1992；Gatev 等人，2006年），以及加入波动率或止损控制机制的扩展形式（Harvey 等人，2018年）。数据驱动型智能体采用现代序列模型：用于长周期预测的 TCN 和 N-BEATS（Oreshkin 等人，2019年），用于高效注意力的 Informer 和 FEDformer（Zhou 等人，2021年；2022年），以及用于零样本迁移的基础模型风格模型 TimeGPT 和 Chronos（Garza 等人，2023年；Ansari 等人，2024年）。新闻与多模态智能体整合文本、价格和情绪，涵盖用于金融语言的 FinBERT（Huang 等人，2023年）以及 FinMME 等多模态基准（Luo 等人，2025年）。基于大语言模型的智能体通过推理和工具使用扩展交易逻辑，代表性实例包括 FinCon、FinMem 和 FinGPT（Yu 等人，2024年；2025年；Liu 等人，2023年），以及将预测模型与大语言模型推理相结合的混合系统（Gan 等人，2025年）。

这些进展展现了金融智能体如何演变为适应性强、数据丰富的决策者。然而，大多数评估仍聚焦于个体表现，而非种群层面的适应，这促使我们采用生态学视角，研究异质智能体如何在共享市场中相互作用并共同适应。

A.2 强化学习

强化学习（RL）为我们在 FinEvo 中的生态学建模提供了互补视角。两者都将市场视为不确定性下的序贯决策过程：强化学习形式化了针对状态—动作轨迹的策略优化（Sutton 等人，1998年），而 FinEvo 则将策略群体的动态描述为随机复制子—突变子过程。两者的相似之处显而易见：强化学习中最大化长周期回报（Mnih 等人，2015年；Schulman 等人，2017年），与演化博弈中策略的生存和主导性相呼应；不过，强化学习侧重于单一自适应策略，而 FinEvo 研究的是异质群体。

强化学习借鉴了多种传统。动态规划和最优控制启发了基于价值的方法（Bellman，1966；Puterman，2014年）；多臂老虎机和在线学习强调遗憾最小化（Auer 等人，2002年；Bubeck 等人，2012）；控制即推断则将规划与概率建模联系起来（Levine，2018年）。这些抽象与 FinEvo 中收益驱动的选择、创新和环境扰动建模相呼应。
方法论的进展进一步拉近了这两个领域的联系。A3C（Mnih 等人，2016年）等可扩展的行动者–评论家算法，凸显了并行性和样本效率在训练自适应策略中的作用。离线强化学习和离策略评估（Kumar 等人，2020年；Uehara 等人，2020年）体现了我们对反事实条件下可复现评估的关注。风险敏感强化学习（Mihatsch 和 Neuneier，2002年）与我们将冲击视为生态压力因素的处理方式相契合。偏好与反馈驱动的强化学习（Christiano 等人，2017年）则类似于我们的创新机制，即外部信号会改变演化轨迹。与此同时，决策 Transformer（Chen 等人，2021年）等序列决策形式展示了监督学习与强化学习如何汇聚，这与 FinEvo 将预测与适应相结合的做法相似。

从这一视角看，强化学习和 FinEvo 可被视为同一范式的两种实现：在不确定性下进行学习并应对延迟反馈，同时平衡利用、探索与鲁棒性。然而，强化学习通常优化的是孤立策略，因此异质策略如何在市场压力下共同演化，仍是一个悬而未决的问题。FinEvo 通过将强化学习原理扩展至智能体群体，弥合了个体层面优化与生态动态之间的鸿沟。

A.3 生态博弈

生态博弈论通过建模策略群体的适应与生存，而非理性个体之间的均衡，扩展了经典博弈论（Hofbauer & Sigmund, 1998年；Sandholm, 2010年）。这类理论源于生物学，其中复制子方程等动态机制（Taylor & Jonker, 1978年）刻画了成功策略如何传播，而突变与漂变则带来持续不断的创新。

这一视角通过进化算法（B¨ack 等，1997年；Deb 等，2002年）、多智能体进化博弈（Tuyls 等，2006年）和基于种群的训练（Jaderberg 等，2017年）影响了机器学习。近期研究强调开放式演化，例如 AlphaStar 的联赛训练（Vinyals 等，2019年）、用于探索的自适应课程（Leibo 等，2019年）以及开放式策略生态（Grill 等，2020年）。这些研究都指出，稳健性源于多样化种群中的持续适应。

在金融领域，生态模型解释了异质交易者之间的交互（Lux & Marchesi，1999年；Hommes，2006年），并在基于智能体的市场中得到实例化（LeBaron，2006年；Farmer & Foley，2009年）。近年来，ABIDES 等高保真平台（Byrd 等，2020年；Amrouni 等，2021年）为市场微观结构研究提供了基础设施，但缺乏关于种群动态的形式化保证。与 FinEvo 思路更接近的是 HFTE 模型（Mahdavi-Damghani，2017年），它为高频量化策略提出了一个多物种捕食者–猎物生态系统，但依赖固定拓扑和遗传算法，而非随机复制子–突变动力学。FinEvo 通过提供一种显式的演化博弈论随机微分方程，将异质策略群体与连续双向拍卖微观结构联系起来，从而补充了这条研究路线。

FinEvo 不止于提供基础设施，还引入了一套有原则的生态形式体系及随机复制子–突变动力学，从而能够对稳健性和适应进行可复现且基于理论的评估。

维度	ABIDES	Mesa	PyMarketSim	TraderTalk	FinEvo（我们的）
目标 / 指标	市场微观结构 结构 仿真	通用 基于智能体的 建模	强化学习基准 用于交易 策略	由大语言模型驱动的 双边交易 行为	**机制 层面评估 演化中的金融 生态系统 tems**
理论保证	None	None	有限（仅训练稳定性）	None	**单纯形不变性、正性、** 奥恩斯坦–乌伦贝克过程 **收敛性**
机制分层	仅智能体级	仅智能体级	智能体策略 交互	对话驱动的 行动	**选择 / 创新- 新 / 扰动- 分解-**
研究用途	压力测试特- 定微观结构- ture	通用智能体 动态演示	强化学习策略 基准测试	大语言模型角色 扮演 交易员	**持仓 系统性评估 压力测试 test, 机制 归因**
可复现性 取决于	配置	高（但非 财务关注点）	中等（对随机种子敏感 （相对于随机种 子）	低（大语言模型变 异性）	**高，具有 跨随机种子鲁棒性 健壮性和 理论一 致性**

表 6：FinEvo 与现有基于智能体的金融市场平台的比较。

B 从 FinEvo 随机微分方程到离散仿真更新

本附录将我们模拟器中实现的离散时间更新规则从连续时间*FinEvoSDE*推导出来。目标是明确理论动力学如何映射为实际更新。

B.1 连续时间表述

设 $X_t = (x_1(t) \dots, x_K(t))^\top \in \Delta^{K-1}$ 为人口占比，其中 $\sum_k x_k(t) = 1$。FinEvo 动力学为

$$dX_t = \underbrace{\beta\, \mathrm{diag}(X_t)\big(V(t) - \bar{V}(t)\mathbf{1}\big)}_{\text{Selection}}\, dt + \underbrace{\mu\,(m_t - X_t)}_{\text{Innovation}}\, dt + \underbrace{\gamma\, \mathrm{diag}(X_t)\, \sigma\, P(X_t)\, dW_t}_{\text{Perturbation}}, \quad (1)$$

其中 $\bar{V}(t) = X_t^\top V(t)$，$m_t \in \Delta^{K-1}$ 是创新分布，$\sigma = \mathrm{diag}(\sigma_1, \dots, \sigma_K)$，而 $P(X) = I - \mathbf{1}X^\top$ 将扰动投影到单纯形切空间上。

B.2 步骤 1：选择–扰动流

首先考虑不含创新项的方程 1。对于每个坐标，定义 $y_k = \log x_k$。根据伊藤引理，在短时间间隔内，若 $V(t)$ 固定，则其一阶近似为：

$$dy_k \;\approx\; \beta\,(V_k - \bar{V})\,dt + \gamma\,\sigma_k\,dB_k,$$

其中 B_k 为相互独立的布朗运动（投影后）。对 Δt 积分可得

$$y_k(t+\Delta t) - y_k(t) \;\approx\; \beta\,(V_k - \bar{V})\,\Delta t + \gamma\,\sigma_k\,\sqrt{\Delta t}\,\xi_k, \quad \xi_k \sim \mathcal{N}(0,1). \quad (2)$$

对方程 2 两边取指数，得到乘法权重更新

$$\hat{x}_k(t+\Delta t) = \frac{x_k(t)\,\exp\big(\beta V_k(t)\Delta t + \gamma\sigma_k\sqrt{\Delta t}\,\xi_k\big)}{\sum_j x_j(t)\,\exp\big(\beta V_j(t)\Delta t + \gamma\sigma_j\sqrt{\Delta t}\,\xi_j\big)}. \quad (3)$$

归一化确保 $\hat{} X_{t+\Delta t} \in \Delta^{K-1}$。

B.3 第 2 步：注入创新

在方程 1 中重新引入创新项后，前向欧拉步给出

$$X_{t+\Delta t} = (1 - \mu\Delta t)\,\hat{X}_{t+\Delta t} + \mu\Delta t\, m_t. \quad (4)$$

因此，种群是选择–扰动更新与创新分布的凸组合。

B.4 最终离散更新

仿真时，我们取 $\Delta t = 1$，得到

$$x_k^{t+1} = (1-\mu)\,\frac{x_k^t\,\exp\big(\beta V_k^t + \gamma\sigma_k\xi_k^t\big)}{\sum_j x_j^t\,\exp\big(\beta V_j^t + \gamma\sigma_j\xi_j^t\big)} + \mu\, m_k^t, \qquad \xi_k^t \sim \mathcal{N}(0,1). \quad (5)$$

实际使用的更新公式是公式 5。通过构造，它保持了正性和单纯形。

B.5 收益动态

每个收益 V_k 都按照奥恩斯坦–乌伦贝克（OU）过程演化，

$$dV_k(t) = \lambda_k(\bar{f}_k - V_k(t))\,dt + \nu_k\,dB_k(t), \quad (6)$$

其精确离散化为

$$V_k^{t+1} = V_k^t e^{-\lambda_k\Delta t} + \bar{f}_k(1 - e^{-\lambda_k\Delta t}) + \nu_k\sqrt{\tfrac{1-e^{-2\lambda_k\Delta t}}{2\lambda_k}}\,\zeta_k^t, \quad \zeta_k^t \sim \mathcal{N}(0,1). \quad (7)$$

B.6 总结

- 选择 $\to$ 按收益进行指数加权 V_k，- 扰动 $\to$ 通过 $\sigma_k\xi_k^t$施加对数正态冲击，- 创新 $\to$ 与 m_t进行凸混合，- 收益 V_k 遵循公式 7 的 OU离散化。

方程5与方程7共同给出了与连续时间 FinEvoSDE 相一致的离散时间模拟器。

C 机制分解

我们旨在计算各策略扰动强度 σ_k 与全局尺度 γ 下扰动项的二阶矩。采用正文中使用的切向（去均值）构造，逐分量离散扰动为

$$\mathrm{N}_k^t \;=\; \gamma\, x_k^t \Big(\sigma_k\, \eta_k^t \;-\; \sum_j x_j^t \sigma_j\, \eta_j^t\Big), \qquad \eta_k^t \sim \mathcal{N}(0,1) \text{ i.i.d. across } k,t.$$

我们计算

$$\mathbb{E}\big[(\mathrm{N}_k^t)^2\big] = \gamma^2 (x_k^t)^2\, \mathbb{E}\left[\Big(\sigma_k\, \eta_k^t \;-\; \sum_j x_j^t \sigma_j\, \eta_j^t\Big)^2\right].$$

展开平方项，并利用独立性和单位方差 $\mathbb{E}[\eta_i^t \eta_j^t] = \delta_{ij}$，可得

$$\begin{aligned}\mathbb{E}\left[\Big(\sigma_k\, \eta_k^t \;-\; \sum_j x_j^t \sigma_j\, \eta_j^t\Big)^2\right] &= \sigma_k^2 \;-\; 2\,\sigma_k \sum_j x_j^t \sigma_j\, \mathbb{E}[\eta_k^t \eta_j^t] \;+\; \sum_{j,\ell} x_j^t x_\ell^t \sigma_j \sigma_\ell\, \mathbb{E}[\eta_j^t \eta_\ell^t] \\ &= \sigma_k^2 \;-\; 2\, x_k^t \sigma_k^2 \;+\; \sum_j (x_j^t)^2 \sigma_j^2.\end{aligned}$$

因此，

$$\mathbb{E}\big[(\mathrm{N}_k^t)^2\big] = \gamma^2 (x_k^t)^2 \left(\sigma_k^2 - 2 x_k^t \sigma_k^2 + \sum_j (x_j^t)^2 \sigma_j^2\right). \tag{8}$$

这个针对每种策略的表达式取代了统一的-σ 公式 $\sigma^2(x_k^t)^2\big(1 - 2x_k^t + \|x^t\|_2^2\big)$。令所有 $\sigma_k \equiv \sigma$ 和 $\gamma \equiv 1$，即可得到统一情形，即 $\mathbb{E}\big[(\mathrm{N}_k^t)^2\big] = \sigma^2(x_k^t)^2\big(1 - 2x_k^t + \|x^t\|_2^2\big)$.

C.1 独立性与方差可加性。

为证明机制分析中所用的方差分解合理，我们将离散更新分解为围绕确定性漂移的三个中心化（零均值）随机波动分量：

$$\Delta x_k^t = \underbrace{S_k^t}_{\text{selection fluct.}} + \underbrace{I_k^t}_{\text{innovation fluct.}} + \underbrace{N_k^t}_{\text{perturbation fluct.}} + \text{(deterministic drift)}.$$

只有中心化波动会影响短期方差。我们在给定 X_t 的条件下进行分析，并作出以下假设。

假设。

(A1)（选择波动）将收益信号写为 $V_k^t = \bar{V}_k^t + \zeta_k^t$，其中 $\bar{V}_k^t := \mathbb{E}[V_k^t \mid X_t]$ 且 ζ_k^t 为中心化波动。定义

$$S_k^t \;:=\; \beta x_k^t \big(\zeta_k^t - \bar{\zeta}^t\big), \qquad \bar{\zeta}^t := \sum_j x_j^t \zeta_j^t,$$

使得 $\mathbb{E}[S_k^t \mid X_t] = 0$。我们假设 ζ^t 具有对角条件协方差 $\mathrm{Cov}(\zeta_k^t, \zeta_j^t \mid X_t) = \delta_{kj}\, \nu_k^2/(2\lambda_k)$。

(A2)（创新波动）给定 X_t，$m_t \sim$ 狄利克雷分布(α) 独立于（ζ^t，η^t），其均值为$m := \mathbb{E}[m_t]$，中心化波动为 $\tilde{m}_t := m_t - m$。定义

$$I_k^t \;:=\; \mu\, \tilde{m}_k^t,$$

so that $\mathbb{E}[I_k^t \mid X_t] = 0$ and $\mathrm{Var}(I_k^t \mid X_t) = \mu^2\, \frac{m_k(1-m_k)}{\alpha_0+1}$.

(A3)（扰动波动）环境冲击 $\eta^t = (\eta_1^t, \ldots\ldots, \eta_K^t)^\top$ 独立同分布于标准正态分布，并且给定 X_t 时独立于（ζ^t, m_t）。定义

$$N_k^t \;:=\; \gamma\, x_k^t \Big(\sigma_k \eta_k^t - \sum_j x_j^t \sigma_j \eta_j^t\Big),$$

使得 $\mathbb{E}[N_k^t \mid X_t] = 0$，而 $\mathrm{Var}(N_k^t \mid X_t)$ 由式 (8) 给出。

引理（方差可加性）。在 (A1)–(A3) 条件下，

$$\mathrm{Var}(\Delta x_k^t \mid X_t) = \mathrm{Var}(S_k^t \mid X_t) + \mathrm{Var}(I_k^t \mid X_t) + \mathrm{Var}(N_k^t \mid X_t).$$

证明。根据构造，$\mathbb{E}[S_k^t \mid X_t] = \mathbb{E}[I_k^t \mid X_t] = \mathbb{E}[N_k^t \mid X_t] = 0$。(A2)–(A3) 中的独立性意味着交叉协方差为零：$\mathrm{Cov}(S_k^t,\ I_k^t \mid X_t) = 0$，$\mathrm{Cov}(S_k^t,\ N_k^t \mid X_t) = 0$，$\mathrm{Cov}(I_k^t,\ N_k^t \mid X_t) = 0$。因此，

$$\mathrm{Var}(\Delta x_k^t \mid X_t) = \mathrm{Var}(S_k^t + I_k^t + N_k^t \mid X_t) = \sum_{\bullet \in \{S,I,N\}} \mathrm{Var}(\bullet_k^t \mid X_t). \qquad \square$$

备注。(i) 确定性漂移（例如 $\beta x_k^t(\bar{V}_k^t - \bar{V}^t)$ 和 $\mu(m_k - x_k^t)$）会影响条件均值，但不会影响中心化方差。(ii) 若要允许收益冲击与环境冲击相关，可将 (A1)–(A3) 替换为具有分块协方差的联合高斯分布；此时方差将包含显式的交叉项 $\mathrm{Cov}(S_k^t,\ N_k^t \mid X_t)$ 等。我们的实证结果采用 (A1)–(A3)，这是演化／复制子随机微分方程中的标准假设，并可导出正文中给出的闭式表达式。

C.2 参数汇总。

符号	描述	典型用途/角色
x_k^t	在时刻 t 使用 π_k 的智能体比例	$\sum_k x_k^t = 1$
f_k^t	策略 k 在 t 的经验平均收益	观测适应度
β	选择强度	调节参数
γ_{kj}	j 对 k的生态效应	基线中设为 0；>0竞争
σ_k	策略 k 的扰动强度	校准或固定
η_k^t	标准正态噪声	模拟抽样
μ	创新概率	调节参数
m_k	创新分布	用户定义
Δt	时间步（通常为 1）	模拟轮次

C.3 结论。

这一推导表明，我们在仿真中实现的智能体级消亡、选择与替换机制，是对加入创新机制的广义 Lotka–Volterra 随机生态学的一种数值上一致的离散化。由此形成的统一框架既具有演化动力学的理论基础，也适用于复杂的基于智能体的市场研究。

D 相关收益与环境冲击下的方差分解

本节将附录 C.1 中的方差分析扩展到收益冲击与环境冲击相关的情形。我们强调，相关性不会 改变 FinEvo 随机微分方程本身；只有分解会增加协方差项。

FinEvo 随机微分方程

$$dX_t = \underbrace{\beta\, X_t \odot (V_t - \bar{V}_t)}_{\text{Selection}}\, dt \;+\; \underbrace{\mu(m_t - X_t)}_{\text{Innovation}}\, dt \;+\; \underbrace{\sigma\, P(X_t)\, dB_t}_{\text{Perturbation}}, \tag{C.1}$$

其中，$P(X_t)$ 是确保单纯形不变性的切空间投影，$B_t \in \mathbb{R}^K$ 是向量值布朗运动。

D.1 允许相关冲击

现在我们假设布朗增量满足

$$\mathrm{Cov}(dB_t) \;=\; \Sigma\, dt, \tag{9}$$

其中， Σ 是一个对称半正定相关矩阵：

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1K} \\ \rho_{21} & 1 & \cdots & \rho_{2K} \\ \vdots & & \ddots & \vdots \\ \rho_{K1} & \cdots & \rho_{K(K-1)} & 1 \end{pmatrix}.$$

这一放宽仅影响方差分解，不影响随机微分方程动力学。

D.2 单个分量的方差

对于任意策略总体分量 $X_{k,\ t}$，伊藤引理给出：

$$\begin{aligned} \mathrm{Var}(dX_{k,t}) = \underbrace{\mathrm{Var}\big[\beta X_{k,t}(V_{k,t} - \bar{V}_t)\big]}_{\text{Selection}} \;+\; \underbrace{\mathrm{Var}[\mu(m_{k,t} - X_{k,t})]}_{\text{Innovation}} \\ +\; \underbrace{\sigma^2 \left(P\Sigma P^\top\right)_{kk}}_{\text{Perturbation}} \;+\; \underbrace{2\sigma\, \mathrm{Cov}\big(\beta X_{k,t}(V_{k,t} - \bar{V}_t),\ (P\, dB_t)_k\big)}_{\text{Selection–Perturbation Cross Term}}. \end{aligned} \tag{C.2}$$

D.3 完整矩阵方差

对于完整向量过程 X_t，我们有：

$$\begin{aligned} \mathrm{Var}(dX_t) = \beta^2\, \mathrm{Var}\big[X_t \odot (V_t - \bar{V}_t)\big] \;+\; \mu^2\, \mathrm{Var}(m_t - X_t) \\ +\; \sigma^2\, P\Sigma P^\top \;+\; 2\beta\sigma\, \mathrm{Cov}\big(X_t \odot (V_t - \bar{V}_t),\ P\, dB_t\big). \end{aligned} \tag{C.3}$$

D.4 与非相关情形的关系

在附录 C.1 中的独立性假设 (A1–A3) 下，我们有

$$\Sigma = I, \qquad \mathrm{Cov}\big(\text{selection},\ dB_t\big) = 0,$$

且方程 (C.2)–(C.3) 化简为正文中的简洁加性分解。当冲击相关时，FinEvo 随机微分方程本身保持不变，方差分解中仅出现额外的协方差项，这与附录 C.1 中的说明一致。

E 最小理论保证

E.1 连续时间 FinEvo 随机微分方程（供参考）。

我们明确写出离散更新背后的连续时间动力学：

$$\begin{aligned} dX_i = X_i \left(f_i(X,t) - \sum_j X_j f_j(X,t) \right) dt \\ +\, \mu\,(m_i - X_i)\, dt \\ +\, \sigma_i X_i \left(dW_i - \sum_j X_j\, dW_j \right), \qquad i = 1, \ldots, K. \end{aligned} \tag{10}$$

其中 $X(t) \in \Delta^K := \{x \in \mathbb{R}^K_{\geq 0} : \sum_i x_i = 1\}$ 为策略份额， $m \in \mathrm{int}(\Delta^K)$，且 $\min_i m_i > 0$，$\mu \geq 0$，而 $\{W_i\}$为相互独立的布朗运动。该扩散与 Δ^K 相切。

E.2 假设。

(i) $f_i(\cdot, t)$ 关于 X 局部满足利普希茨条件，且具有线性增长界；该条件在紧时间区间上一致成立；(ii) $m \in \text{int}(\Delta^K)$, $\mu \geq 0, \sigma_i \geq 0$; (iii) 布朗运动 $\{W_i\}$ 相互独立。与外生价值过程 $V_i(t)$ 耦合时，我们假设 V_i 为具有唯一平稳分布且矩有界的奥恩斯坦–乌伦贝克过程；离散奥恩斯坦–乌伦贝克近似见方程 7。

引理1（适定性）在上述假设下，对于任意 $X(0) \in \Delta^K$，随机微分方程 *10* 在 $[0, \infty)$ 上存在唯一的强解。

命题1（单纯形不变性与正性）设 $X(0) \in \Delta^K$。则方程 *10* 的解满足 $\sum_i X_i(t) \equiv 1$，对所有 $t \geq 0$ 和 $X_i(t) \geq 0$ 几乎必然成立。此外，若 $X(0) \in \text{int}(\Delta^K)$、$\mu > 0$，且 $\min_i m_i > 0$，则对所有 $t > 0$，$X_i(t) > 0$ 几乎必然成立（边界是非吸收性的）。

证明概要。将方程 10 对 i 求和，选择漂移项和切向扩散项均相互抵消；创新项的和为 $\mu(\sum_i m_i - \sum_i X_i) = 0$，因此质量守恒。乘性扩散在边界处消失，这意味着非负性。当 $\mu > 0$ 且 $\min_i m_i > 0$ 时，漂变 $\mu(m_i - X_i)$ 严格指向内侧。□

命题2（不变测度与准稳定性）考虑联合马尔可夫过程$(X(t)\ V(t))$，其中 X 由方程*10*控制，V为奥恩斯坦–乌伦贝克过程。若 $\mu > 0$ 且 $\min_i m_i > 0$，则在 $\Delta^K \times \mathbb{R}^K$上至少存在一个不变概率测度。此外，在小扰动／弱创新状态下（$\sigma \to 0$，$\mu \to 0$满足 $\mu = O(\sigma^2)$），不变测度集中于确定性复制子–突变子常微分方程 $\dot{X}_i = X_i\,(f_i - \sum_j X_j f_j) + \mu(m_i - X_i)$的吸引子集附近，从而产生准稳定的生态组成。

证明概要。奥恩斯坦–乌伦贝克过程具有唯一的平稳分布，且矩有界。根据命题1，$X(t)$ 始终处于紧单纯形内；联合过程是费勒过程且具有紧性，因此克雷洛夫–博戈柳博夫定理保证不变测度的存在性。小扰动下的集中性可由耦合系统的标准弗雷德林–温策尔论证得到。□

E.3 离散时间对应形式（一致性）。

我们的模拟更新先执行指数选择，再进行创新/突变。选择归一化步骤为式 equation 3，创新混合为最终更新式 equation 5。二者共同定义了一个保持正性的单纯形不变映射。具体而言，

$$\sum_k x_k^{t+1} = 1, \quad x_k^{t+1} \geq 0, \quad x_k^{t+1} > 0 \text{ for } t \geq 1 \text{ if } \mu > 0 \text{ and } \min_k m_k > 0.$$

耦合价值信号采用式 equation 7 中的OU离散化。在 $\Delta t \to 0$ 下，该离散格式是式 equation 10 的一阶相容近似。

F 智能体库与初始化。

F.1 智能体名称映射

表 F.1 总结了全文所用智能体全称与其缩写之间的映射。智能体按类别组织（基于规则、信息型、强化学习、时间序列模型，以及基于大语言模型的交易者）。该映射提供了简明参照，以确保正文中的图、表与分析保持一致。

F.2 智能体库

本节介绍我们框架中每种智能体原型的完整设计，包括其交易目标、信息来源、决策规则和执行机制。通过详细介绍经典策略（如技术型交易者和信息型交易者）以及现代自适应智能体（如深度学习、强化学习和基于大语言模型的模型），智能体库展现了 FinEvo 生态模拟中各种行为的异质性。

表 7：智能体名称映射

类别	全称	缩写
技术交易者	动量交易者	MT
	RSI交易者	RSI
	突破交易者	BT
知情交易者	保守型知情交易者	CIT
	激进知情交易者	AIT
	价差感知知情交易者	SAIT
	缩放知情交易者	SIT
	机会主义知情交易者	OIT
强化学习	Double DQN交易者	Double DQN
	PPO交易者	PPO
	A2C交易员	A2C
时间序列模型	iTransformer交易员	ITT
	Informer交易员	IT
	TimeMixer交易员	TimeMixer
基于大语言模型的智能体	LLM ReAct 交易员	ReAct
	LLM 思维链交易员	CoT
	大语言模型交易交易者	交易
	LLM FinCon 交易员	FinCon

F.2.1 信息型交易代理

- **交易目标**
  利用市场价格 P_t 与预言机提供的基本价值 P_t^*之间的差异，在不产生长期方向性偏差的情况下进行盈利交易。
- **信号计算**
  在每个时刻 t，计算相对偏差

$$\delta_t = \frac{P_t - P_t^*}{P_t^*}, \quad P_t = \frac{P_t^b + P_t^a}{2}$$

- **决策规则**
  给定偏差阈值 $\theta > 0$，智能体的交易决策 $a_t^i \in$ {买入，卖出，持有} 定义如下：

$$a_t^i = \begin{cases} \text{buy}, & \text{if } \delta_t < -\theta \text{ and } h_t^i \leq 0 \\ \text{sell}, & \text{if } \delta_t > \theta \text{ and } h_t^i > 0 \\ \text{hold}, & \text{otherwise} \end{cases}$$

  其中 h_t^i 表示智能体 i在时间 t的当前库存头寸。
- **执行机制**
  根据不同变体，智能体按以下方式提交订单：
  – 保守型：下达**限价单**，价格为

$$P^{\text{buy}} = P_t^b + \epsilon, \quad P^{\text{sell}} = P_t^a - \epsilon$$

  – 激进型：立即以最优卖价/买价**下**市价单。
  – 机会型：下**市价单**，数量为

$$q_t = \begin{cases} \lambda\, q, & \text{if liquidity is low} \\ q, & \text{otherwise} \end{cases}$$

  其中， q 是基础交易数量， $\lambda > 1$ 是放大因子。

F.2.2 做市商代理

- **交易目标**
  Maintain liquidity 通过在多个价格层级持续提交买卖双方的限价单，在限价订单簿中进行交易，旨在从买卖价差中获利。
- **信号计算**
  在每个决策时刻 t，观察最优买价和卖价：
  $$P_t^b, \quad P_t^a$$
  计算中间价和半价差：
  $$P_t^m = \frac{P_t^b + P_t^a}{2}, \quad s_t = \frac{P_t^a - P_t^b}{2}$$
  从预定义区间中为每一侧抽取订单总量：
  $$q_t \sim [q_{\min}, q_{\max}]$$
  随机确定两侧各自报价的价格档位数量 L。
- **决策规则**
  对于每个档位 $i = 0$，1，.....，$L-1$，计算限价单价格：
  $$P_{t,i}^b = P_t^m - s_t - i, \quad P_{t,i}^a = P_t^m + s_t + i$$
  根据权重向量 $w = (w_1$，.....，$w_L)$分配各档位的订单量，满足
  $$\sum_{i=1}^{L} w_i = 1, \quad q_{t,i} = w_i \cdot q_t$$
- **执行机制**

  – **轮询模式（subscribe=False）**：查询当前市场价差，并围绕中间价以对称限价单挂单，每个档位的订单规模相同。– **订阅模式（subscribe=True）**：订阅订单簿更新，随机选择 $L \in \{1, \ldots, 5\}$ 个档位，并根据预定义的权重向量 w从档位报价字典中分配交易量。

F.2.3 噪声交易者代理

- **交易目标**
  无法获取基本面信息或技术指标。通过随机决策来模拟非理性或无信息交易行为。

- **时间 t的决策规则**
  该智能体以相同概率 $p = 0.5$ 执行以下两种动作之一：
  – *Agg*激进成交（市价单）：
  以相同概率提交市价买入或卖出订单：
  $$q_t \sim \mathcal{U}\{10, 20, ..., 100\}, \quad a_t = \begin{cases} \text{MarketBuy}(q_t), & \text{prob. } 0.5 \\ \text{MarketSell}(q_t), & \text{prob. } 0.5 \end{cases}$$
  – 被动执行（限价单）：
  提交一笔价格与中间价存在偏差的限价买入或卖出订单：
  $$P_t^m = \frac{P_t^b + P_t^a}{2}, \quad \epsilon_t \sim \mathcal{U}(0.001, 0.005)$$
  $$P_t^{\text{limit}} = \begin{cases} \lfloor P_t^m (1 - \epsilon_t) \rfloor, & \text{if buy} \\ \lceil P_t^m (1 + \epsilon_t) \rceil, & \text{if sell} \end{cases}$$
  $$a_t = \begin{cases} \text{LimitBuy}(q_t, P_t^{\text{limit}}), & \text{prob. } 0.5 \\ \text{LimitSell}(q_t, P_t^{\text{limit}}), & \text{prob. } 0.5 \end{cases}$$
- **下一决策时间**
  下一个唤醒时间从以下分布中采样：
  $$t_{\text{next}} = t + \Delta t, \quad \Delta t \sim \mathcal{U}(55, 65) \text{ seconds}$$

F.2.4 事件驱动的新闻反应智能体

- **交易目标**
  React 通过根据新闻的情绪方向立即下市价单，响应突发市场新闻事件（例如宏观经济冲击或特定资产评级）。
- **信息来源**
  订阅一个事件总线，该总线会广播离散的新闻事件。每个事件包括：
  - 事件类型 $\in$ {正面、负面、上调、下调，..... }
  - 内容：文本描述（逻辑会忽略）
- **事件接收时的决策逻辑**
  在时间 t，收到新闻事件 e_t 时：

$$\text{Action}_t = \begin{cases} \text{MarketBuy}(q), & \text{if } \texttt{event\_type} \in \{\texttt{POSITIVE}, \texttt{UPGRADE}\} \\ \text{MarketSell}(q), & \text{if } \texttt{event\_type} \in \{\texttt{NEGATIVE}, \texttt{DOWNGRADE}\} \\ \text{NoAction}, & \text{otherwise} \end{cases}$$

  其中 q 为固定订单数量（例如 $q = 10$ 股）。
- **时序与唤醒频率**
  Th 智能体 10 秒唤醒一次，以维持内核活动并确保只订阅一次事件流，但交易决策完全由事件驱动。
- **成交机制**
  所有交易都会在信号接收后立即以**市价单** 提交，以确保对重大事件做出最快响应。

F.2.5 强化学习交易者 AGENT

- **交易目标**
  基于强化学习的智能体旨在通过以在线方式与环境交互，最大化累计投资组合收益。在每个时间步，它都会观察当前市场状态，选择一个动作（买入/卖出/持有），并通过持续学习（在线微调）更新其策略参数。
- **状态表示**Th
  e mar kket state s_t is represented as follows for different learning schemes:

$$\begin{aligned} &\textbf{Q-Learning:} && s_t = (x_t^{(1)},\ x_t^{(2)},\ x_t^{(3)}) \text{ (discretized)} \\ &\textbf{PPO:} && s_t = \left[x_t^{(1)},\ x_t^{(2)},\ \ldots,\ x_t^{(n)}\right] \text{ (continuous and normalized)} \end{aligned}$$

  1
- **动作策略**
  Th智能体选1取离散空间中的动作：

$$\mathcal{A} = \{0 = \text{HOLD},\ 1 = \text{BUY},\ 2 = \text{SELL}\}$$

  每个动作都会按对应方向触发数量为 Q 的市价单。
- **奖励函数**
  智能体根据投资组合价值的变化获得奖励：

$$r_t = PV_t - PV_{t-1} - \lambda \cdot \mathbb{I}(a_t \in \{\text{BUY}, \text{SELL}\})$$

  其中，PV_t 为投资组合价值，λ 为交易惩罚系数。
- **学习机制**

  1. **Q学习（表格型、在线）**
     智能体维护一个Q表（λ，μ），并使用标准的贝尔曼方程进行更新：

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)\right]$$

     参数（α，γ）分别表示学习率和折扣因子。根据最近的经验缓冲区 D，每经过 N 步在线进行更新。

2. **近端策略优化（PPO）**
The 智能体 l获得随机 hastic poli性 $\pi_\theta(a|s)$，使用裁剪代理目标：

$$L^{\text{PPO}} = \mathbb{E}_t \left[\min \left(r_t(\theta) A_t,\ \text{clip}(r_t(\theta), 1-\epsilon, 1+\epsilon) A_t \right) \right]$$

其中 $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}$ 是概率比，A_t 是优势函数。智能体收集 N 个交互步数，并执行 K 轮梯度更新以进行在线微调。

F.2.6 TECHNICALTRADER 智能体

- **交易目标**
  根据历史价格数据推导出的常见技术分析信号执行市价单，旨在捕捉趋势或均值回归机会。

- **状态观测**
  在每个决策时刻 t，智能体都会观测近期价格窗口：

$$S_t = \{P_{t-n}, \ldots, P_t\}$$

  其中，P_t 是最新的成交价格。

- **交易数量**Fi
  xed order size Q 对于所有交易。

- **交易策略与决策规则**
  智能体采用以下策略之一：

  1. 动量智能体（移动平均线交叉）
     计算短期和长期移动平均线：

$$MA_t^{short} = \frac{1}{N_s} \sum_{i=t-N_s+1}^{t} P_i, \quad MA_t^{long} = \frac{1}{N_l} \sum_{i=t-N_l+1}^{t} P_i$$

     交易操作 a_t 为：

$$a_t = \begin{cases} \text{BUY}, & \text{if } MA_t^{short} > MA_t^{long} \text{ and position} \leq 0 \\ \text{SELL}, & \text{if } MA_t^{short} < MA_t^{long} \text{ and position} \geq 0 \\ \text{HOLD}, & \text{otherwise} \end{cases}$$

     按 a_t方向下达数量为 Q的市价单。
  2. *RSI*智能体（均值回归）
     计算相对强弱指数（RSI）：

$$RSI_t = 100 - \frac{100}{1 + RS}, \quad RS = \frac{\text{Average Gain}}{\text{Average Loss}}$$

     交易操作 a_t 为：

$$a_t = \begin{cases} \text{BUY}, & RSI_t < \theta_{\text{low}} \\ \text{SELL}, & RSI_t > \theta_{\text{high}} \\ \text{HOLD}, & \text{otherwise} \end{cases}$$

     按 a_t 方向下达数量为 Q 的市价单（RSI低时买入，RSI高时卖出）。
  3. 突破智能体（通道突破）
     定义突破水平：

$$R_t = \max(P_{t-L}, \ldots, P_{t-1}), \quad S_t = \min(P_{t-L}, \ldots, P_{t-1})$$

     交易操作 a_t 为：

$$a_t = \begin{cases} \text{BUY}, & P_t > R_t \\ \text{SELL}, & P_t < S_t \\ \text{HOLD}, & \text{otherwise} \end{cases}$$

     随后下达数量为 Q的市价单，a_t。

F.2.7 深度学习智能体

- **交易目标**
  Predict the 未来 th
  使用基于市场指标训练的深度神经网络预测资产价格的走向，并据此做出交易决策。

- **输入与预测**
  在每个决策时刻 t，智能体都会接收一个经过处理的市场指标向量（例如相对强弱指数、平滑异同移动平均线、价差等），记作 x_t。智能体使用参数化神经网络$f_\theta(x_t)$，生成对应于三类的逻辑值：价格下跌、不变或上涨。
  $$\hat{y}_t = f_\theta(x_t) \in \mathbb{R}^3, \quad \text{with action}_t = \arg\max_i \hat{y}_{t,i}$$

- **决策规则**
  根据预测类别：
  – 如果动作$_t =$ 为上涨：下达数量为 Q的市价买单– 如果动作$_t =$ 为下跌：下达数量为 Q的市价卖单– 否则：持有

- **在线监督学习**
  Th 智能体在线 收集 i (x_t, y_t) coll 经验元组 y_t，其中 y_t 是经过固定预测时域后已实现的未来价格方向。这些数据会存储在训练缓冲区中，并定期通过随机梯度下降微调模型参数 θ：
  $$\theta \leftarrow \theta - \eta \cdot \nabla_\theta \mathcal{L}(f_\theta(x), y)$$
  其中 $\mathcal{L}$ 是交叉熵损失。

- **学习周期**
  智能体在以下步骤之间交替进行：
  1. **预测：** 使用当前模型 f_θ 做出交易决策 2. **标注：** 等待预测时域结束，为旧的观测值添加标签
  3. **微调：** 使用缓冲区中近期的已标注样本训练 f_θ

F.2.8 TIMEMIXERTRADING 智能体

- **交易目标**
  使用深度学习模型（TimeMixer）预测交易资产的短期价格方向，并据此下达市价单。模型会利用新观测到的数据持续进行在线微调，以提升预测性能。

- **信息来源**
  Periodically $DataCenterAgent$ 请求标准化市场指标向量。这些指标可能包括动量、流动性、订单簿失衡和归一化持仓等度量。

- **接收指标时的决策逻辑**
  在时间 t，收到指标向量 x_t 后，TimeMixer 模型会输出交易信号 $\hat{y}_t \in \{$买入，卖出、持有$\}$，并采取以下动作：
  – $\hat{y}_t =$ 卖出：预测价格下跌 $\Rightarrow$ 下达市价卖出(q)– $\hat{y}_t =$ 持有：预测价格没有显著变化 $\Rightarrow$ 不采取行动– $\hat{y}_t =$ 买入：预测价格上涨 $\Rightarrow$ 下达市价买入(q)
  其中 q 表示固定的订单数量。

- **标注与在线训练**
  每条观测 x_t 都会在预测时域 ΔT 之后，根据实际价格变动进行存储和标注。如果相对变化超过上阈值 θ_1，标签为“上涨”；如果低于下阈值 θ_2，标签为“下跌”；否则标签为“中性”。经验（x_t，y_t）会存储在回放缓冲区中。智能体会定期从缓冲区中采样小批量，并通过随机梯度下降微调模型参数 θ 。

- **时序与唤醒频率**
  Th智能体在每个仿真步骤中被唤醒，并可在等待响应期间安排短期唤醒以维持活动。

- **执行机制**
  所有交易均作为**市价单** 以固定规模提交，以确保立即成交。智能体根据已执行交易动态更新其现金与持仓。

F.2.9 iTransformer交易智能体

- **交易目标**
  用于监督其短期价格预测的iTransformer模型。该智能体利用预测信号提交市价单，同时使用流式市场数据持续微调模型。

- **信息来源**
  定期向*DataCenterAgent*请求技术指标向量（例如动量、订单簿不平衡、流动性价差、持仓特征）。这些指标经过归一化并转换为iTransformer的输入特征。

- **接收指标时的决策逻辑**
  收到指标向量 x_t后，iTransformer 会输出一个交易信号 $\hat{y}_t \in${买入，卖出，持有}:

  – $\hat{y}_t =$ 买入 $\Rightarrow$ placeMarketBuy(q)– $\hat{}$
  $y_t =$ 卖出 $\Rightarrow$ place MarketSell(q)– $\hat{}$
  $y_t =$ 持有 $\Rightarrow$ 执行不操作

  其中 q 表示固定的交易数量。

- **标注与在线训练**
  每个观测都会在预测时域 ΔT之后，根据实际价格变动相对于过去水平的情况分配一个标签。如果变动超过上阈值 θ_1，标签为“买入”；如果低于下阈值 θ_2，标签为“卖出”；否则标签为“持有”。这些已标注样本会存储在回放缓冲区中。定期抽取小批量样本，并使用随机梯度下降微调模型参数。

- **时序与唤醒频率**
  智能体按随机间隔安排唤醒，以请求新的指标。在等待期间，使用短期唤醒来重试或维持活动。

- **执行机制**
  所有交易均以固定规模的**市价单** 执行，以确保立即成交。每笔交易后，现金和持仓组合都会动态更新。

F.2.10 信息型交易者智能体

- **交易目标**
  交易者将市场价格与基本价值（由预言机提供）进行比较。如果相对偏差超过预定义阈值，他们便会根据各自的特定策略下单。

- **信息来源**
  Th智能体从预言机提取基本价值，并从*ExchangeAgent*维护的订单簿中获取最佳买卖报价。根据这些信息，计算偏差比率：

$$d_t = \frac{p_t^{\text{market}} - p_t^{\text{fundamental}}}{p_t^{\text{fundamental}}}.$$

- **决策逻辑**
  基类（InformedTraderAgent）提供了计算偏差和提交订单的通用机制。子类会重写 take action() 方法，以实现不同的交易风格：

– **保守型信息交易代理**当偏差超过阈值时下达**限价单**，旨在进行被动交易。当市场低估资产时买入；当市场高估资产时卖出，但仅限于当前持仓允许的情况下。– **激进型信息交易代理**一旦偏差超过阈值，便立即通过**市价单**作出反应。当资产被低估时执行买入，当资产被高估时执行卖出。– **机会型信息交易代理**

在交易前评估可用流动性。若订单簿对手方的流动性不足，则扩大订单规模；否则使用标准数量。订单以市价单提交。

– **规模调整型信息交易代理**
根据偏离幅度按比例调整订单规模。偏离越大，交易数量越大，直至上限倍数。交易使用限价单下达。

– **价差感知型信息交易代理**
根据买卖价差在限价单与市价单之间进行选择。若价差较宽，则优先使用限价单以获取更优价格；若价差较窄，则以市价单立即成交。

- **时序与唤醒频率**
该智能体会在市场交易时段内以随机间隔唤醒，重新评估偏离情况、更新持仓，并在适当时提交订单。
- **执行机制**
执行方式取决于子类：被动型（限价单）、激进型（市价单）或混合型（感知价差或按比例调整）。所有订单消息都会通过交易所路由，并记录下来以进行状态跟踪。

F.2.11 TECHNICALTRADER 智能体

- **交易目标**
技术交易者依据历史价格模式和衍生出的技术指标做出交易决策。他们旨在根据各自的策略捕捉短期价格趋势、反转或突破走势。

- **信息来源**
每个智能体都会观察市场近期的交易价格，并计算移动平均线、相对强弱指数或价格通道水平等技术指标。这些指标用于判断市场趋势或超买/超卖状态。

- **决策逻辑**
这些智能体采用不同的交易风格：
  – **动量智能体**
  采用双移动平均线交叉策略。短期平均线向上穿越长期平均线（上升趋势）时买入，短期平均线向下穿越长期平均线（下降趋势）时卖出。订单以市价单的形式下达。

  – **RSI智能体**
  使用相对强弱指数（RSI）进行均值回归交易。当市场处于超卖状态（RSI 低于阈值）时买入，处于超买状态（RSI 高于阈值）时卖出。该智能体还使用市价单进行成交。

  – **突破智能体**
  采用通道突破策略。若当前价格超过近期高点（阻力位），智能体就买入；若跌破近期低点（支撑位），则卖出。订单使用市价单立即成交。

- **时序与唤醒频率**
智能体以随机间隔醒来，观察市场、计算指标并评估交易机会。随机性使智能体之间产生去同步化，避免人为制造的同时性。

- **成交机制**
  所有智能体均根据各自指标得出的信号下达市价单。交易执行会被记录，持仓也会相应更新。智能体将现金、股票持仓和投资组合价值作为其内部状态的一部分进行跟踪。

F.2.12 基于大语言模型的 TradingAgents

大语言模型整合市场指标与新闻信息，并运用不同的推理机制生成交易动作。

- **LLM ReAct 交易员 交易目标：**将市场指标与新闻文本纳入 ReAct 推理过程，以生成即时交易动作，同时记录可解释的推理。**状态观测：**在每个决策时刻 t，智能体观测处理后的市场指标向量 x_t 以及近期新闻 n_t。**决策规则：**智能体执行 ReAct 推理（推理 $\rightarrow$ 行动），以选择动作$a_t \in$ {买入、卖出、持有}：根据 a_t 下达固定数量 Q 的市价单。**日志记录：**推理步骤与中间分析会被存储，以供可解释性使用。

$$a_t = \text{ReAct}(x_t, n_t)$$

- **LLM 思维链交易员 交易目标：**利用指标与新闻生成思维链（思维链）推理过程，以提高决策质量。**状态观测：**在时刻 t，观测 x_t 与 n_t。**决策规则：**构建一系列推理步骤：

$$\text{CoT}_t = \{s_1, s_2, \ldots, s_k\}, \quad a_t = \text{FinalDecision}(\text{CoT}_t)$$

数量为 Q 的市价单根据最终动作执行。

- **LLMTrading 交易员交易目标：**根据指标和新闻信息生成结构化交易计划（动作 + 数量）。**状态观测：**在决策时刻 t观测 x_t 和 n_t 。**决策规则：**生成交易计划：

$$\text{Plan}_t = \{\text{action}_t, \text{quantity}_t\} = \text{LLMTradingModel}(x_t, n_t)$$

智能体按照拟定的计划执行市价单。

- **LLM FinCon 交易员交易目标：**在金融约束（例如最大持仓限额）下生成交易动作，并综合考虑市场指标和新闻。**状态观测：**在时刻 t，观测 x_t、 n_t和当前持仓 h_t。**决策规则：**计算受约束的动作：

$$a_t, q_t = \text{ConstrainedDecision}(x_t, n_t, h_t, Q, h_{\max})$$

仅当市价单符合金融约束时才会下单。推理和计划数量会被记录。

G 实验详情与结果

G.1 人工冲击下的仿真

交易价格左侧面板展示了在13:00注入向上冲击后的平均价格轨迹，右侧面板则展示了在12:00注入向下冲击后的动态变化。在这两种情况下，蓝线表示模拟交易价格的平均路径，阴影区域表示各次运行结果的变异范围，即均值上下的95%置信区间。向上冲击导致价格持续上涨，而向下冲击则引发价格骤跌，随后逐渐下滑，这与典型市场反应十分相似。这些结果进一步证实，我们的框架能够在极端外生扰动下产生合理的市场响应，从而增强了所提出的生态动力学模型的有效性。



图10：模拟外生冲击对交易价格的影响。蓝线表示平均交易价格路径，阴影区域表示不同仿真运行中的 $\pm 3\sigma$ 置信带。

冲击诱导的网络动态

作为对主要分析的补充，我们进一步考察外部冲击如何重塑策略之间的结构性交互。图11和图12展示了正向冲击下共现矩阵和互信息网络的演变，而图13和图14则展示了负向冲击下的相应结果。这些视角提供了关于稳健性的补充证据，并揭示了联盟如何在冲击下重新组织。

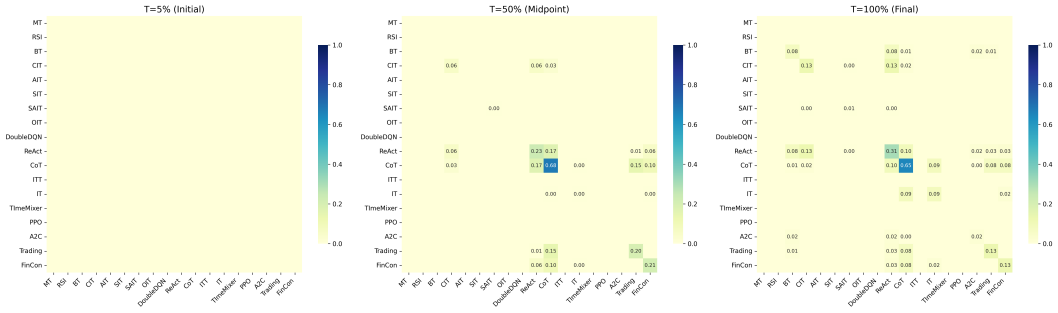


图11：正向冲击下策略共现矩阵的演化。数值表示在不同阶段（初期、中期、末期），策略对超过主导阈值的频率。

日内网络演化（2025年7月7日）

为考察正常交易日（无外生冲击）内生态结构如何演化，我们追踪 2025年7月7日主导策略的共现 情况，以及策略之间的互信息（*MI*）网络。时间按交易时段归一化（T=0%，25%，50%，75%，100%）。在共现视图中，条目（i，j）记录策略 i和 j 同时占主导地位（份额为 $>$ 12%）的频率。在MI网络中，节点大小表示策略的主导性，边的粗细则表示共依赖强度。

详细分析。（早期， T=0% – 25%）共现几乎可以忽略，MI网络较为稀疏，表明这是一个弱协调的碎片化生态。大多数策略彼此独立地运作。

（午间， T=50%）共现矩阵中首次出现共同主导对，MI网络变得枢纽中心化：一两个策略充当连接者，将原本互不相连的集群连接起来。这标志着市场状态从探索行为转向协调性专业化。

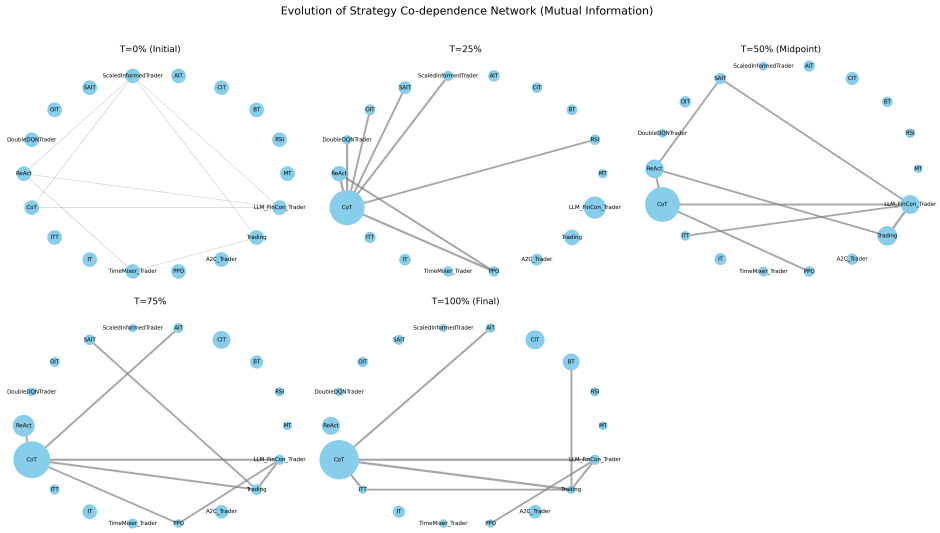


图12：正向冲击下互信息网络的演化。节点大小反映策略主导性，边的粗细表示共依赖强度。

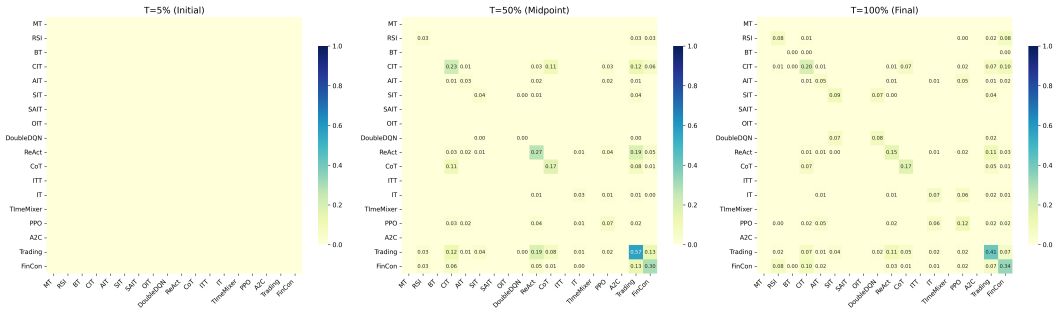


Figure 13：负向冲击下策略共现矩阵的演化。与 th 相比 e positive 情形下，网络更快地重新集中，并出现更少但更强的联盟 .

（较晚时段， $T=75\%-100\%$）少数联盟得以持续并不断强化，而边缘策略则逐渐脱钩。MI网络呈现选择性剪枝：多余的连接逐渐消失，高权重联系集中于主要枢纽周围。这与日终群体占比重新集中的现象相一致。

要点。 在一天之中，生态从分散（低共现、稀疏互信息）走向有组织（稳定联盟、枢纽中心化互信息），随后 趋于稳定 直至收盘。这种日内重组与我们的宏观发现一致：探索在早期占主导，而选择则逐步将联盟巩固为一个小型的持久骨干。

G.2 跨多个真实市场交易日的额外日内轨迹 S

在正文（图6）中，我们在两个较长的真实市场时期评估了 FinEvo：*2022*年*1*月至*10*月（熊市状态）和 *2024*年*10*月–*2025*年*7*月（牛市状态）。这些实验已经在**多个月**的真实数据上汇总了生态系统统计数据。然而，由于篇幅限制，正文中仅展示了一条具有代表性的日内轨迹。

因此，我们提供了**涵盖牛市和熊市多个日期的日内群体演化轨迹**。这些图表表明，FinEvo 的生态动态并非只适用于某一个交易日：在各种交易条件下，主导地位变化、策略更替和阶段变化模式均保持一致。

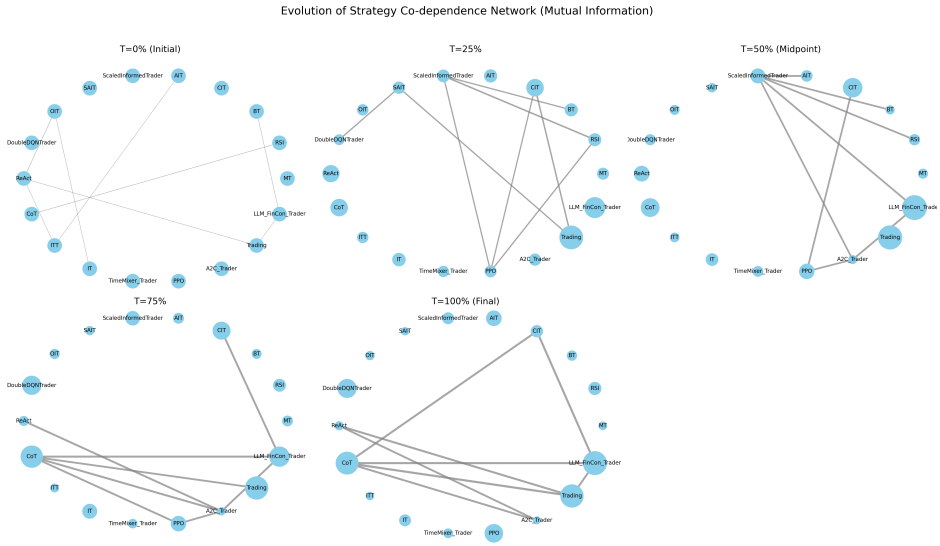


图 14：负向冲击下互信息网络的演化。结果凸显了多元联盟的瓦解，以及由大语言模型智能体主导的集中式枢纽的兴起。

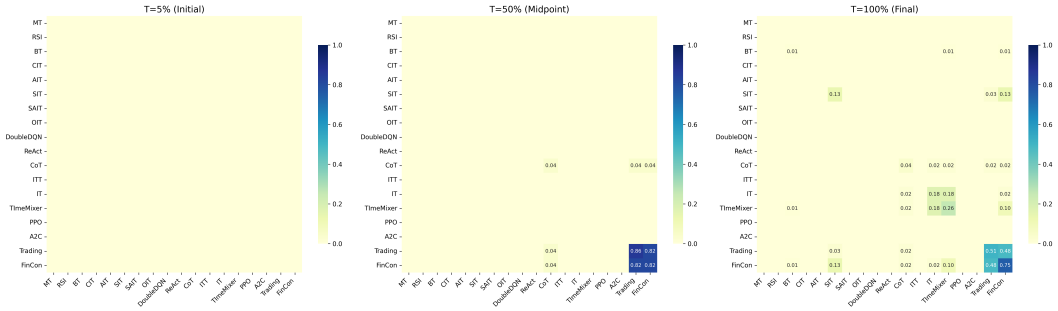


图 15：2025年7月7日**共现**矩阵的日内演化（主导阈值>12%）。日初（T=0%）矩阵几乎为空；日中（T=50%）少数配对开始共同占据主导；临近收盘时（T=100%），结构逐渐围绕少数持久联盟巩固，而许多策略仍处于边缘。

这些结果证实，FinEvo 在真实数据上的表现**在不同交易日、月份和市场状态下均具有稳健性**，

G.3 对初始智能体参数扰动的稳健性

为评估 FinEvo 的生态动态是否对初始智能体级超参数高度敏感，我们开展了一项参数扰动稳健性实验。对于每次蒙特卡洛运行，所有关键超参数（例如强化学习学习率、基于规则的智能体的窗口长度，以及大语言模型提示参数）均围绕其基线值随机扰动 ±20%，并从均匀分布中采样。这样，每次仿真都会产生显著不同的异质初始化。

在牛市和熊市状态下，扰动参数版本在所有典型事实指标上都与基线高度吻合。这证实，厚尾、下行不对称性、夏普比率离散度和波动率聚集是**FinEvo 演化动力学中稳定涌现的属性**，而非特定超参数选择造成的假象。

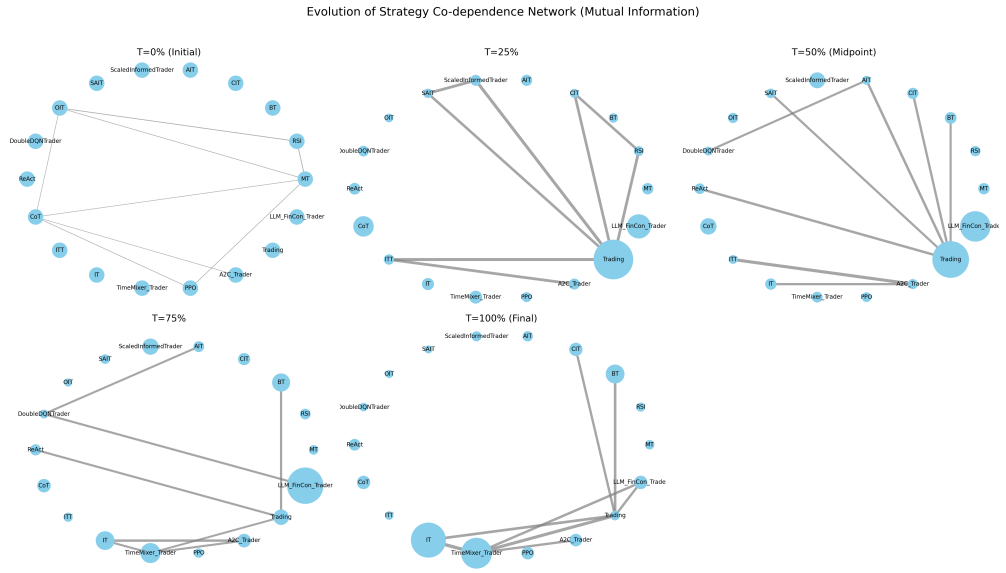


图 16：2025年7月7日互信息网络的日内演化。节点大小表示主导性；边的粗细表示共依赖强度。网络从早间稀疏、弱连接的构型，转变为午间的枢纽中心结构，随后经过选择性剪枝并逐渐稳定，直至收盘。

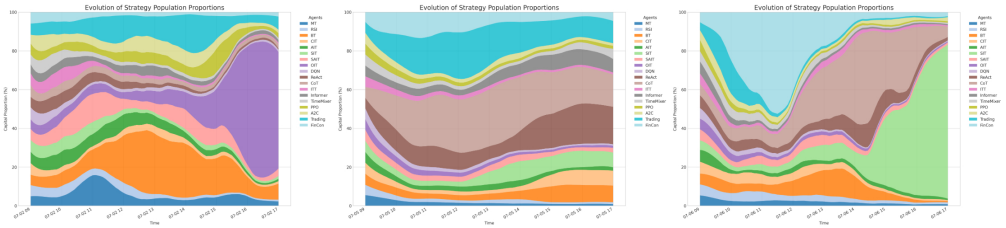


图 17：**2025年7月2日、07-05日和07-06日的日内种群演化。**

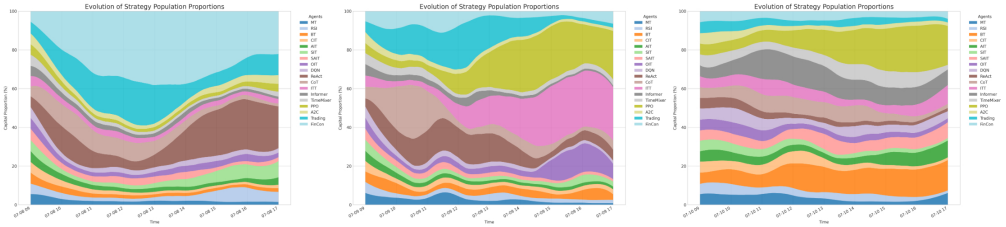


图 18：**日内　种群演化，日期为2025年7月8日、07-　　09日和07-10日。**

表 8：受扰动智能体参数下的稳健性（128 次运行）。数值报告均值 ± 95%置信区间。典型事实指标在 ±20%扰动下保持稳定。

市场状态	模型	超额峰度$_{\text{mean}}$↑	偏度$_{\text{mean}}$↓	夏普比率离散度$_{\text{mean}}$↑	波动率的波动率$_{\text{mean}}$↑
看涨	FinEvo	3.163 ± 0.138	−1.532 ± 0.129	1.669 ± 0.061	0.008 ± 0.0004
	FinEvo（扰动的）	3.141 ± 0.152	−1.596 ± 0.138	1.638 ± 0.067	0.008 ± 0.0004
看跌	FinEvo	4.259 ± 0.172	−1.793 ± 0.149	1.732 ± 0.089	0.013 ± 0.0007
	FinEvo（扰动的）	4.196 ± 0.181	−1.692 ± 0.157	1.684 ± 0.094	0.012 ± 0.0008

G.4 对相关重尾冲击的稳健性

在正文中，我们的分析结果是在独立高斯冲击这一标准假设下推导得出的。为验证我们的实证发现并非这种简化带来的假象，我们开展了额外的稳健性实验，其中收益冲击和环境冲击均具有相关性且呈重尾分布。具体而言，我们对驱动价值过程 $V_k(t)$的噪声进行如下修改。在每个时间步 t，我们不再抽取相互独立的标准正态增量，而是构造策略层面冲击

$$\epsilon_k(t) \;=\; \rho\, Z_{\text{common}}(t) \;+\; \sqrt{1-\rho^2}\, Z_{k,\text{idio}}(t), \tag{11}$$

其中，$\rho \in (0,\ 1)$控制相关强度（在我们的实验中取 $\rho = 0.5$），$Z_{\text{common}}(t)$是所有策略共享的共同宏观因子，而$Z_{k,\ \text{idio}}(t)$是策略 k特有的特异性成分。$Z_{\text{common}}(t)$和$Z_{k,\ \text{idio}}(t)$均取自自由度为 ν 的标准化 Student-t 分布（我们使用 $\nu = 5$），并经过缩放以达到单位方差。由此得到策略层面正相关的重尾冲击：

$$\text{Cov}\big(\epsilon_k(t), \epsilon_\ell(t)\big) \;=\; \rho^2 \quad (k \neq \ell), \tag{12}$$

与主要分析中的独立高斯设定形成对比。随后，我们用这些相关重尾冲击驱动每种策略的价值过程：

$$dV_k(t) \;=\; \lambda_k\big(m_t - V_k(t)\big)\, dt \;+\; \nu_k\, \epsilon_k(t)\, \sqrt{dt}, \tag{13}$$

对于环境冲击，我们取

$$\eta_t \;=\; Z_{\text{common}}(t), \tag{14}$$

因此，策略层面的收益明确地彼此相关，也与环境冲击相关。我们采用与第3节相同的市场配置，在这一相关重尾设定下重新运行了主要模拟。所得市场仍表现出：(i) 肥尾且负偏的损益分布，(ii) 显著的波动率聚集和更高的波动率的波动，以及 (iii) 策略群体中类似的生态相变。从定量角度看，典型事实指标仅发生中等幅度的变化；详细比较见表 9。这些结果证实，独立性／正态性假设主要用于随机微分方程推导中的解析上的可处理性，对 FinEvo 的定性行为并非至关重要。

Ta表 9：对相关重尾冲击的稳健性（128 次运行）。数值报告均值 ± 95% 置信区间 .

市场状态	模型	超额峰度$_{\text{mean}}\uparrow$	偏度$_{\text{mean}}\downarrow$	夏普比率离散度$_{\text{mean}}\uparrow$	波动率的波动率$_{\text{mean}}\uparrow$
看涨	FinEvo	3.163 ± 0.138	-1.532 ± 0.129	1.669 ± 0.061	0.008 ± 0.0004
	FinEvo（相关）	5.278 ± 0.151	-1.632 ± 0.133	1.727 ± 0.075	0.011 ± 0.0007
看跌	FinEvo	4.259 ± 0.172	-1.793 ± 0.149	1.732 ± 0.089	0.013 ± 0.0007
	FinEvo（相关）	6.399 ± 0.214	-1.932 ± 0.175	1.844 ± 0.101	0.015 ± 0.0009

G.5 跨资产稳健性

设置。 FinEvo 的设计不依赖特定资产：演化随机微分方程和 CDA机制均不基于任何资产特定假设。为考察其在异质市场中的稳健性，我们在四种广泛研究的股票——AAPL、TSLA、MSFT 和 XOM——上实例化相同的 FinEvo 生态。在所有资产中，演化参数（β, μ, γ）、智能体库、价值学习动力学和市场微观结构均完全相同。唯一依赖资产的组成部分是经验波动率尺度，以及用于参数化外生冲击的资产特定新闻时间表。因此，每种资产代表的都是不同经验环境下的同一演化系统，而非单独校准的模型。

结果。 表10报告了策略损益分布四项典型事实指标的均值和95%置信区间。在所有资产中，FinEvo 均持续呈现正超额峰度（肥尾）、负偏度（下行不对称性）、较高的波动率的波动率（波动率聚集），以及显著的夏普比率离散度（异质策略表现）。这些指标的幅度也呈现出符合经济直觉的差异：TSLA 的厚尾、偏度、夏普比率离散度和波动率的波动率最为突出；AAPL 处于中间范围；MSFT 和 XOM 则表现出较温和的尾部特征和较低的状态切换强度。这些发现表明，无需针对单一资产进行调优，FinEvo 的演化动力学在异质资产间仍具有稳健性。所有资产均采用与第3.3节相同的经验熊市时期（2022年1月至10月）进行评估，因此，跨资产差异仅源于波动率尺度和新闻模式，而非市场状态不同。

Tab表 10：FinEvo 的跨资产稳健性。各项指标报告为 128 次运行的均值 $\pm$ 95% 置信区间 s.

资产	超额峰度	偏度	夏普比率离散度	波动率的波动率
AAPL	4.354 ± 0.41	-1.656 ± 0.18	1.633 ± 0.12	0.011 ± 0.0015
TSLA	$\mathbf{5.578 \pm 0.53}$	$\mathbf{-2.344 \pm 0.24}$	$\mathbf{1.975 \pm 0.18}$	$\mathbf{0.015 \pm 0.0023}$
MSFT	3.176 ± 0.35	-0.967 ± 0.15	1.158 ± 0.09	0.008 ± 0.0016
XOM	3.741 ± 0.38	-0.573 ± 0.12	1.392 ± 0.17	0.008 ± 0.0011

在我们的实验中，对于 AAPL、TSLA、MSFT 和 XOM，相同的 FinEvo 配置会产生定性上相似的生态特征（表 10），这表明该框架并不严格依赖于某一种特定的底层资产。虽然我们在此仅测试股票，但进化随机微分方程和 CDA机制并不依赖于股票特有的假设，因此原则上，采用适当的数据和校准后，同一形式体系可应用于其他电子交易市场（例如期货、外汇、加密货币）。

G.6 实验配置

价格撮合机制（连续双向拍卖）

订单簿与优先级。我们采用价格—时间优先的连续双向拍卖。在时间 t，最佳买价/卖价分别为 b_t 和 a_t，价差为 $s_t = a_t - b_t$。买卖双方在价位 ℓ 的深度分别为 $B_t(\ell)$ 和 $A_t(\ell)$，最小变动价位为 δp。同一价格的订单按先进先出顺序排队。

订单类型。智能体提交限价单(p, q)（正数 q 表示买入，负数表示卖出）或市价单$(p = \pm\infty)$。如果限价单 $p \geq a_t$ （买入）或 $p \leq b_t$ （卖出），则该订单可立即成交。

撮合规则与成交。到达的可立即成交数量 $q > 0$ （买入）会消耗卖盘阶梯$\{(A_t(1)\ a_t),\ (A_t(2)\ a_t + \delta p)$ $\ldots\}$，直到全部成交或订单簿耗尽。如果订单扫过 L 个价位，各价位的交易量为 v_ℓ，价格为 p_ℓ，则成交量加权执行价格（VWAP）为

$$\overline{p}_t \;=\; \frac{\sum_{\ell=1}^{L} v_\ell\, p_\ell}{\sum_{\ell=1}^{L} v_\ell}, \qquad \sum_{\ell=1}^{L} v_\ell \;=\; q_{\text{filled}} \leq q.$$

任何未成交的剩余数量 $\Delta q \;=\; q - q_{\text{filled}}$ （对于限价单）都会以其限价 p 挂在相应一侧，时间戳为 t。最后成交价格 P_t 设为该事件中最后一笔撮合的价格；随后订单簿更新深度和最优报价。

中间价／微观价格与滑点。定义中间价 $m_t = \frac{a_t + b_t}{2}$和深度加权微观价格$\tilde{m}_t = \frac{A_t(1)\, b_t + B_t(1)\, a_t}{A_t\ 1\ + B_t\ 1\ ()()\ 。}$对于买入，瞬时滑点（相对于交易前中间价）为

$$\text{Slip}_t^{\text{buy}} \;=\; \overline{p}_t - m_{t^-}, \qquad \text{Slip}_t^{\text{sell}} \;=\; m_{t^-} - \overline{p}_t.$$

交易摩擦（成本）。我们通过每笔交易按比例收取 c =0.5% 的成本（包括滑点/冲击设定），来反映现实中的摩擦。对于名义金额为 $N_t = |p_t\, q_{\text{filled}}|$的一笔交易，费用部分为 $c\, N_t$。

现金、库存与按市值计价。对于智能体 k，设库存为 $x_{k,t}$ ，现金为 $C_{k,t}$。一笔成交的<size $q > 0$ 的买入交易，价格为 p_t ，将更新

$$C_{k,t} \leftarrow C_{k,t} - \overline{p}_t\, q - c\,|\overline{p}_t q|, \qquad x_{k,t} \leftarrow x_{k,t^-} + q.$$

对称地，卖出产生 $C_{k,\ t} \leftarrow C_{k,\ t} + p_t\,|q| - c\,|p_t q|, \quad x_{k,\ t} \leftarrow x_{k,\ t^-} - |q|$。事件时点的按市值计价投资组合价值 t为

$$V_{k,t} \ = \ C_{k,t} + x_{k,t}\, m_t.$$

损益与收益率。增量损益为 Δ损益$_{k,\ t} = V_{k,\ t} - V_{k,\ t^-}$。单个交易时段的累计收益率为

$$\mathrm{Ret}_k \ = \ \frac{V_{k,T} - V_{k,0}}{V_{k,0}}.$$

我们报告 Ret_k 以及跨次运行的 95% 置信区间（CI）：$\mathrm{CI}_{95} = \bar{r} \pm 1.96\hat{}\sigma/\sqrt{n}$。 $^-$

夏普比率、回撤、换手率。使用事件级或 K 线级收益率 $\{r_{k,t}\}$，年化夏普比率为

$$\mathrm{Sharpe}_k \ = \ \frac{\mu(r_{k,t})}{\sigma(r_{k,t})}\,\sqrt{A},$$

其中 A 为年化因子（例如，每年的 K 线数）。最大回撤为 $\mathrm{MDD}_k \ = \max_{t \le u}\big(1 - V_{k,\ u}$ 最大值$_{s \le t} V_{k,\ s}\big)$。交易时段内的换手率按交易名义金额与平均权益之比计算：

$$\mathrm{Turnover}_k \ = \ \frac{\sum_t |\overline{p}_t\, q_{k,t}|}{\frac{1}{T}\sum_t V_{k,t}}.$$

解读。

- 价格—时间优先 确保公平排队：更优价格优先成交；在同一价格档位内，较早的时间戳优先于较晚的时间戳。
- *VWAP*执行 刻画多档位消耗，并纳入相对于 m_t–的微观结构滑点。
- *0*每笔交易*.5%*的成本惩罚过度交易，并使盈利能力更接近现实条件。
- 按市值计价损益 在中间价通过 m_t 将库存风险与同期流动性联系起来（或在需要时使用$\tilde{m}_t$ ）。
- 夏普比率/最大回撤/换手率 共同刻画风险调整后绩效、路径风险和交易强度，作为原始收益的补充。

G.6.1 实验设置

所有实验均使用模拟市场引擎进行，该引擎提供了一个具有真实订单簿动态的连续双向拍卖市场（价格撮合机制的详情请参见附录 G.6）。我们通过扩展环境，加入三个关键组件来捕捉生态演化动态：

1. **新闻发布线路**。 外部信息冲击通过新闻发布渠道注入，该渠道会在预定时间向所有智能体发送共同信号。冲击可以是正面的，也可以是负面的，并直接影响智能体信念和交易激进程度。2. **多指标分配机制**。 仿真不仅进行简单的价格更新，还会根据多项指标追踪并重新分配资本，这些指标包括已实现损益、经波动率调整的收益以及策略多样性。这使绩效压力、创新和环境扰动能够共同影响生态适应。3. **演化规则**。 我们实现了第 2 节介绍的复制子-突变子动态。在每个时期（设为仿真时间 5 分钟），智能体种群占比会根据选择强度 β、突变率 μ和扰动强度 σ进行更新。随后，资本将按照更新后的比例重新分配，以确保生态真实性并维持市场价值守恒。

除非另有说明，每种配置均使用 128 个蒙特卡洛随机种子运行，每天的交易时段为 8 小时。交易成本（包括滑点）设为每笔交易 0.5%。所有智能体的初始资本相等，并在同一个合成市场环境中交互。

G.6.2 智能体参数

为清晰起见，我们总结了实验中使用的基础设施智能体的设计。由于每种智能体类型的交易逻辑已在附录 F.2.1 中详细说明，此处仅介绍其功能角色和关键配置参数。

DataCenterAgent。 作为唯一的全局数据中心和指标引擎。它从交易所收集交易和流动性快照，计算一系列微观结构和技术指标（例如买卖价差、订单簿失衡、相对强弱指数、平滑异同移动平均线、ATR、布林带），并通过有界双端队列维护历史记录，以确保内存效率。关键参数包括：

- 唤醒频率： 5 秒（更新采样间隔）。
- 最大历史长度：200 个刻度。
- 服务所需的最短历史长度：30 个刻度。
- 输出：indicator history.pkl 和 liquidity history.pkl。

EventBusAgent。实现一个集中的发布/订阅中心。它管理其他智能体的订阅，并转发广播事件（例如新闻）。该设计轻量简洁：订阅者存储在集合中，以确保去重；事件则实时分发给所有已注册的智能体。

NewsAgent。按照预先指定的时间表发布外生事件（例如新闻冲击）。每到预定时间，事件便会通过 EventBusAgent 广播。时间表定义为一个已排序的（t，事件）对列表，以确保可复现性，并精确控制信息冲击的时间。

NoiseTraderAgent。一种基线随机交易者，会向市场注入随机流动性。每次唤醒时，它会以 50% 的概率下达市价单（立即成交）或限价单（以接近中间价的价格下单，并附带小幅随机偏移）。**关键参数：** 交易标的、随机订单规模（10–100 股）、唤醒间隔（55–65 秒，均匀抽取）。

MarketMakerAgent。 在订单簿的买卖两侧对称地提供流动性。每次唤醒时，它都会撤销尚未成交的订单，并在多个价格档位上提交新订单，以覆盖价差。它可采用两种模式运行：(i) 轮询模式（根据当前价差，在中间价附近下单）；(ii) 订阅模式（订阅订单簿更新，并根据预定义的档位报价分布分配交易量）。**关键参数：** 初始现金 = 10M、订单规模范围 = [10, 50]、唤醒频率 = 5s、档位数 = 1–5；订单规模按 DEFAULT LEVELS QUOTE DICT 进行拆分。

NewsReactionAgent。 一种事件驱动交易者，通过事件总线对 NewsAgent 发布的外生新闻作出反应。收到广播后，它会提交与情绪一致的市价单（正面/上调时买入，负面/下调时卖出）。**关键参数：** 初始现金 = 1M、每次反应的交易规模 = 10 股、唤醒频率 = 10s（仅用于维持订阅）。

MomentumAgent. 一种基于移动平均线交叉的趋势跟踪交易者。当短期平均线超过长期平均线时买入，并在情况相反时退出。**关键参数：** 初始现金 = 1M，短期窗口 = 5，长期窗口 = 20，交易规模 = 10，唤醒间隔 = 55–65秒。

RSIAgent。一种使用相对强弱指数（RSI）的均值回归交易者。当 RSI 低于超卖阈值时买入，超过超买阈值时卖出。**关键参数：** 初始现金 = 1M，RSI 窗口 = 14，阈值 = [30, 70]，交易规模 = 10，唤醒间隔 = 55–65 秒。

突破代理。一种通道突破趋势跟随者。当当前价格高于近期最大值时买入，低于近期最小值时卖出。**关键参数：**初始现金 =1M，回溯窗口 = 50，交易规模 = 10，唤醒间隔 = 55–65 秒。

InformedTraderAgent（基类）。这类智能体会根据预言机观察到的市场价格与基本价值之间的偏差进行交易。若偏差超过阈值，则由子类决定执行方式。**关键参数：**初始现金 = 1M，交易规模 = 10，偏差阈值= 2%，唤醒间隔 = 2–3分钟。

- **保守型知情交易者**。当价值偏差较大时，挂出被动限价订单。
- **激进型知情交易者**。检测到偏差后立即执行市价单。
- **机会型知情交易者**。根据可用流动性调整交易规模（例如，卖方深度较薄时加大买入规模）。
- **规模化知情交易者**。根据偏差幅度按比例调整订单规模（最高为基线的3倍）。
- **价差感知型知情交易者**。根据买卖价差的宽度，选择限价单或市价单。

DoubleDQN交易代理。一种在线强化学习交易者，通过经验回放实现 Double DQN。状态空间根据相对强弱指数、订单簿失衡（OBI）和买卖价差分箱进行离散化。该智能体针对三种动作（持有、买入、卖出）学习 Q值，并定期更新目标网络。**关键参数：**学习率 = 10^{-3}，折扣因子 = 0.9, $\epsilon = 0.1$, 缓冲区大小 = 1000，批次 = 32，目标更新间隔 = 50 步，交易规模 = 100，唤醒间隔 = 2–3 分钟。

大语言模型智能体（思维链、ReAct、交易）。我们的通用基于大语言模型的智能体依靠结构化提示，将技术指标、新闻事件和投资组合状态整合起来。它们的推理风格各不相同：思维链采用逐步审慎推演，*ReAct* 将推理与动作交替进行，而交易则采用混合式证据汇总工作流。新闻事件到来时，智能体会立即请求指标并做出交易决策；没有新闻时，它们则以随机间隔（5–15 分钟）唤醒，以避免同步伪影。交易规模按 10 个单位的倍数离散化，从而确保可复现性，并便于在智能体之间进行比较。这些设计选择使智能体能够及时响应外生冲击，同时保持行为的异质性与可处理性；稳健性检查也证实，在采用其他采样间隔和交易粒度时，系统仍具稳定性。

iTransformer交易代理。一种采用监督学习的交易者，使用*iTransformer* 架构。它根据流式指标预测短期价格方向（下跌、中性、上涨），标签通过延迟的价格变动生成。模型通过回放缓冲区进行在线微调。**关键参数：**输入特征 = {相对强弱指数、平滑异同移动平均线、价差、订单簿失衡、持仓}，输出 = 3 动作，缓冲区大小 = 1000，批次 = 32，训练频率 = 10，交易规模 = 10，唤醒间隔 = 2–3 分钟。

Informer交易智能体。一种基于*Informer* 架构、用于时间序列预测的监督学习智能体。它从 DataCenter 获取流式指标，并预测未来价格方向（下跌、持平、上涨）。预测结果用于执行交易，而标签则根据已实现收益事后生成，从而支持借助回放缓冲区进行在线微调。**关键参数：**输入 = {相对强弱指数、平滑异同移动平均线、价差、订单簿失衡、持仓}，预测时域 = 10 分钟，缓冲区 = 1000，批次 = 32，训练频率 =10，交易规模 = 10，唤醒间隔 = 2–3 分钟。

TimeMixer交易智能体。一个使用*TimeMixer* 序列模型的监督式智能体，也采用在线训练。它处理技术特征和持仓状态，预测方向性变动、执行交易，并定期进行微调。**关键参数：**输入 = {相对强弱指数、平滑异同移动平均线、价差、订单簿失衡、持仓}，隐藏维度 = 64，2 层，缓冲区 = 1000，批次 = 32，训练频率 = 10，交易规模 = 10，唤醒间隔= 2–3 分钟。

强化学习智能体（优势演员-评论家、近端策略优化）。我们的强化学习智能体采用金融强化学习文献中常用的经典超参数。例如，折扣因子固定为 $\gamma = 0.99$,

熵系数为 0.01，价值系数为 0.5，符合连续控制基准中的标准做法。训练更新基于较短的 n 步轨迹（$n = 20$），在嘈杂的市场反馈下兼顾响应速度与稳定性。动作空间经过刻意简化，设为 {买入、卖出、持有}，并采用固定交易规模（10 单位），以确保不同智能体之间的可比性。我们还验证了，对这些超参数进行适度扰动不会实质性改变定性动态。

LLM FinCon（风险感知）。 除基于证据的推理外，FinCon还纳入了一个轻量级风险模块。我们使用最近50个已实现收益来近似条件风险价值（CVaR），并在估计的CVaR低于 -2%时触发风险警报。该阈值是根据经验选定的，作为下行脆弱性的保守型代理指标。为避免对单一设定过拟合，我们在不同窗口大小（20–100）和阈值（-1%至-3%）下验证了稳健性，所得定性模式保持一致。当风险警报处于激活状态时，交易规模会被缩小，以体现更为谨慎的行为。

种群管理器。 生态动态由随机复制子–突变机制支配。我们采用演化博弈论标准范围内的参数：选择强度 $\beta = 2.0$、突变率 $\mu = 0.05$以及环境扰动 $\sigma = 0.1$。策略特定的奥恩斯坦–乌伦贝克参数（λ, ν）用于平滑经验收益，类似于生态建模中的均值回归适应度动态。狄利克雷先验 $\alpha = 10$ 鼓励适度的基线多样性。尽管这些取值并未针对特定市场进行校准，但定性行为——如共存、优势循环和灭绝事件——在（$\beta, \mu, \sigma, \alpha$）的适度扰动下仍具稳健性，这一点已由敏感性检验证实。

G.6.3 分析指标

Our 评估框架采用微观–中观–宏观结构，指标定义如下。

微观（个体表现与多样性）。我们跟踪每个智能体的财务表现与生存：

- **收益**：收益$_i = \frac{V_{\text{最终}i} - V_{\text{初始}i}}{V_{\text{初始}i}}$，通过自助法计算95%置信区间。
- **夏普比率**：夏普$_i = \frac{\mathbb{E}[r_{i,t}]}{\sigma(r_{i,t})}$。
- **最大回撤**：最大回撤$_i = \max_{t<u} \frac{P_{i,t} - P_{i,u}}{P_{i,u}}$
- **换手率**：换手率$_i = \frac{\sum_t |q_{i,t}^{buy} + q_{i,t}^{sell}|}{\sum_t q_{i,t}^{pos}}$，用于衡量交易强度。
- **胜率**：智能体获得正收益或高综合得分的运行次数占比。

中观（种群多样性与交互结构）。我们评估中间层级的生态结构：

- **集中度（赫芬达尔-赫希曼指数）**：赫芬达尔-赫希曼指数 $= \sum_k x_k^2$，其中 x_k 是策略 k的人口占比。

- **策略熵**：基于离散智能体动作的分布 $\mathcal{A}$ ={市价买入，限价买入，持有，限价卖出，市价卖出}，我们在每个时间步 t 计算处于各动作状态的智能体占比 p_i（t）。然后计算香农熵

$$H(p(t)) = -\sum_{i=1}^{5} p_i(t)\, \log_2 p_i(t).$$

  较高的 H 表示活动多样且无序，而较低的 H 则表示同质化（例如，羊群效应，所有智能体都执行单一动作）。
- **模块度**：策略相关网络中的基于图的聚类强度。
- **协同作用与对抗**：策略对之间的平均带符号相关性，以热力图呈现。

• **共现与互信息**：策略之间联合主导与依赖结构的频率，反映联盟形成与重构。

宏观（系统波动性与状态转变）。我们将总体波动分解为三个组成部分：

$$\mathrm{Var}(\Delta x_k) = V_{\text{selection}} + V_{\text{innovation}} + V_{\text{perturbation}},$$

分别对应绩效驱动的选择压力、创新与社会混合，以及环境扰动。我们进一步通过**相变**的次数（主导策略转换的频率）来量化体制不稳定性。

相变指标我们给出实证分析中所用相变指标的正式定义。

Let

$$X_t = (x_{1,t}, \ldots, x_{K,t})$$

表示时刻 t 各 K 策略类别的种群份额向量。将主导策略集合定义为

$$\mathcal{D}(t) = \Big\{ k \ : \ x_{k,t} = \max_j x_{j,t} \Big\}.$$

定义。若主导策略集合的成员发生变化，则在时刻 t 发生相变：

$$\boxed{\text{PhaseChange}(t) = 1 \quad \text{iff} \quad \mathcal{D}(t) \neq \mathcal{D}(t^-)}$$

否则，$\text{PhaseChange}(t) = 0$。

分布统计量（系统层面）。令 $\{r_t\}_{t=1}^T$ 表示系统层面收益序列（例如总体市场收益）。我们报告四项在风格化事实文献中常用的关键分布指标。

• **超额峰度（尾部厚重程度）**。实证研究表明，金融收益呈尖峰厚尾且具有显著的非高斯特征：与方差相同的正态分布相比，其分布在尾部和均值附近的概率质量更大（Engle, 1995；Bollerslev, 1986；Engle & Ng, 1993）。超额峰度衡量相对于高斯基准分布的尾部厚重程度。

$$\text{ExKurt}(r) = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_t - \mu_r}{\sigma_r}\right)^4 - 3, \quad \mu_r = \tfrac{1}{T}\textstyle\sum r_t, \quad \sigma_r^2 = \tfrac{1}{T}\textstyle\sum (r_t - \mu_r)^2.$$

高斯分布的超额峰度为 $= 0$；正值表示存在肥尾（与正态模型相比，大幅冲击发生得更频繁），这对应着更高的崩盘风险和收益中的跳跃式行为。

• **偏度（收益分布的不对称性）**。偏度反映大幅正向变动和大幅负向变动是否同样可能发生。许多股票市场呈现负偏度，反映出大幅下行变动（崩盘）的概率高于大幅上行跳跃的概率（Barberis et al., 1998）。

$$\text{Skew}(r) = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_t - \mu_r}{\sigma_r}\right)^3.$$

接近 0 的值对应对称收益；$\text{Skew}(r) < 0$ 表示厚左尾（下行风险占主导），而 $\text{Skew}(r) > 0$ 表示厚右尾。

• **夏普比率离散度（策略表现的异质性）**。超额峰度和偏度是在市场层面计算的，而夏普比率离散度衡量的是同一环境中不同策略类别的表现差异。令 S_k 表示策略类别 k 的夏普比率，$\bar{S} = \frac{1}{K}\sum_{k=1}^{K} S_k$ 表示它们的截面均值。那么

$$\text{SharpeDisp} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}(S_k - \bar{S})^2}.$$

SharpeDisp较低意味着大多数策略获得的风险调整收益相近（环境并未明显区分它们），而SharpeDisp较高则表明该生态系统能鲜明地区分稳健的“赢家”和脆弱的“输家”，从而为交易规则提供严苛的压力测试。

• **波动率的波动率（波动率的不稳定性与聚集性）**。金融市场中的波动率随时间变化，并表现出持续的聚集性：高波动时期往往紧随高波动时期之后出现，平静时期也往往接续平静时期（Andersen等人，2003年）。为捕捉这一特征，我们首先构建波动率的滚动窗口估计量（例如，窗口大小为 W）：

$$\sigma_t = \sqrt{\frac{1}{W}\sum_{i=t-W+1}^{t}(r_i-\mu_t)^2}, \quad \mu_t = \tfrac{1}{W}\sum_{i=t-W+1}^{t} r_i.$$

然后，我们测量这一波动率过程的标准差：

$$\mathrm{VoV} = \sqrt{\frac{1}{T-W}\sum_{t=W}^{T}(\sigma_t-\bar{\sigma})^2}, \qquad \bar{\sigma} = \tfrac{1}{T-W}\sum_{t=W}^{T}\sigma_t.$$

较大的波动率的波动率意味着波动率更加不稳定、聚集性更强（平静与动荡市场状态之间频繁转换），而较小的波动率的波动率则对应更加平稳的同方差市场。

在异质性智能体模型和基于智能体的模型中，公认的评估理念并不是拟合价格轨迹，而是考察模型是否能够内生地产生真实市场中观察到的典型事实（Lux & Marchesi，1999年；LeBaron，2006年；Hommes，2006年）。理论与微观结构研究进一步强调，这些非高斯特征源于市场参与者之间的相互作用，而不是外生噪声或曲线拟合（Bouchaud等人，2009年）。因此，再现典型事实是衡量模型与真实金融行为“接近程度”最有意义且最广泛认可的指标。

经济解释。 该指标刻画的是市场生态的结构性重组，而非一种统计效应。它有别于（i）波动率聚集——一种价格层面的现象，以及（ii）策略区制转换——通常指单个智能体内部的微观层面变化。相变反映的是生态系统均衡在宏观层面的转变，由 FinEvo 随机微分方程的演化力量（选择、创新、扰动）所驱动。

智能体观测到的交易指标。 为确保真实性，智能体依据标准微观结构与技术信号来决定其动作。流动性与流量。 买卖价差、订单簿失衡（OBI）、市场深度（买入/卖出）以及逐笔规则成交流失衡（TFI）。动量与摆动指标。ROC、相对强弱指数、平滑异同移动平均线（线、信号线、柱状图）、Aroon（上行/下行）、KDJ 以及威廉%R。波动率与基于 *OHLC* 的指标。 ATR、布林带（上轨/中轨/下轨）以及 ADX/DMI（ADX、 +DI、 –DI）。这些特征使代理行为立足于市场可观测信息，而非先知信息。

总体而言，这种微观—中观—宏观设计提供了互补的视角：个体的盈利能力与多样性、群体内部的关系结构，以及生态层面的系统性方差分解。三者结合，可对不断演变的市场博弈开展细粒度和总体分析。

G.6.4 人工冲击

我们通过同时扰动以下因素，注入可控的外生冲击：**(i)** 事件驱动/LLM智能体处理的新闻流；**(ii)** 信息交易员观察到的潜在基本面价值过程；并且 **(iii)** 收紧微观结构条件，以模拟流动性压力。除非另有说明，冲击均安排在**12:00–13:30（当地时间）**期间，并报告正面和负面两种情形。

新闻渠道。 设 λ_{base} 为新闻到达的基准泊松率， z_t 为情绪载荷标量。在冲击窗口 $[t_s, t_e]$ 期间，我们同时提高其频率和幅度：

$$\lambda_{\text{shock}} = \kappa\,\lambda_{\text{base}}, \qquad z_t \sim \mathcal{N}(\mu_s, \sigma^2_{\text{news}}),\ \ s \in \{+1, -1\},$$

其中， $s = +1$ 表示正向冲击（$\mu_+ > 0$），$s = -1$ 表示负向冲击（$\mu_- < 0$）。所有订阅新闻总线的智能体（事件驱动型智能体和大语言模型）都会实时接收这些消息；更高频率、更大幅度的信号会立即产生方向性推动。**基本面渠道**。信息交易者会观察由预言机提供的基本面锚点 $P_t^\star$，我们利用它来维持交易价格与现实世界估值之间的稳定联系。冲击是在此锚点之上叠加的，而非取代它。设 $P^\star_{\text{base}}(t)$ 表示基准基本面曲线（冲击发生前）。在新闻冲击开始时 t_s，我们叠加一个与冲击方向一致的确定性隆起，使有效锚点变为

$$\tilde{P}_t^\star \;=\; P^\star_{\text{base}}(t)\,\Big(1 + s\,A_f\,e^{-(t-t_s)/\tau_f}\,\mathbf{1}\{t \ge t_s\}\Big),$$

其中， $s \in \{+1, -1\}$ 表示正向/负向冲击， A_f 控制影响幅度（相对于基线锚点的百分比），而 τ_f 设定衰减期限。随后，信息交易者根据我们实现中的方式，依据相对于˜ ˜的有效锚点的错误定价采取行动：当 $\delta_t = (P_t - P_t^\star)/P_t^\star$时，他们会在 $\delta_t < -\theta$时买入，在 $\delta_t > \theta$时卖出，从而将基本面冲击传递至订单流，同时保持锚点联系。:contentReference[oaicite:1]{index=1}

G.6.5 数据详情

为使日内模拟贴近现实情况，我们构建了一个多源金融数据集，整合了涵盖资产、公司和宏观经济环境的异质信息流。我们特意涵盖两个截然不同的市场状态——2022年1月至10月的全球熊市和2024年10月至2025年7月的牛市，以确保模型在结构迥异的动态环境中具备稳健性。所有信号均按日内分辨率对齐，从而支持高频事件驱动交易和生态适应。

市场事件与新闻。高频事件流来自**GDELT**、**ReutersMarkets**和**彭博终端**，每个交易日以分钟级粒度提供数千条带时间戳的市场相关头条。这些数据构成驱动日内波动率的外生冲击的主要来源。

宏观经济与政策信号。来自**FRED** 的宏观经济指标（例如国内生产总值、消费者价格指数、失业率、美国国债收益率）被建模为预定事件，并在其官方发布时间戳注入。**美联储政策文件**（联邦公开市场委员会会议纪要、声明和新闻发布会）被视为与发布时间对齐的离散日内冲击，提供有关货币政策立场和前瞻指引的文本信号。

公司层面与分析师数据。企业财报、财务报表和**分析师预测**（每股收益、目标价格、行业展望）均按其官方披露时间纳入。由于这些变量变化较慢，因此会将其插值为日频锚点，以为基于基本面的智能体提供信息。

集成与路由。根据各类数据的时间分辨率和结构格式，将其分发给专门的分析智能体：将高频事件流分发给日内新闻智能体，将定期发布的宏观数据分发给冲击敏感型智能体，并将低频基本面数据分发给资产负债表解读智能体。这种路由方式确保异质信号得到联合利用，以支持事件驱动交易和长期适应。

来源	内容	分辨率
GDELT / 路透社 / 彭博社	市场新闻、企业事件	分钟级日内
FRED / 宏观数据发布	国内生产总值、消费者价格指数、失业率、收益率	发布时的日内
公司报告 / 分析师预测	盈利、目标价	季度（每日锚点）
美联储 / 联邦公开市场委员会	政策声明、新闻发布会	发布时的日内

表11：与日内模拟交易时钟相匹配的数据来源和时间分辨率。

G.7 模拟场景的动机与依据

此处我们说明第3节所用两种情景的设计动机与实证依据。

情景1：人工冲击（受控机制验证）。该情景旨在对 FinEvo 的演化动力学进行受控压力测试。我们在一个原本干净的合成市场中注入外生的正向和负向冲击，以此分离并考察扰动如何通过演化随机微分方程中的选择、创新和扰动成分传播。这种受控设置使我们能够清晰追踪冲击放大、生态转变和重新稳定模式；若仅在充满噪声的真实世界条件下进行测试，这些现象将难以解读。此类受控扰动测试是异质智能体模型和系统性风险研究中的标准做法。

情景2：真实世界新闻（实证相关性与稳健性）。在受控环境中验证相关机制后，我们利用高频GDELT/路透社发布内容，在新闻驱动的真实市场条件下评估FinEvo。这检验了该框架在面对非结构化、具有经济意义且会影响短期收益、订单流和交易行为的信号时，能否保持稳定并提供有用信息。该情景还评估异质主体——尤其是基于大语言模型的交易者——能否从真实新闻中提取有用信号，并在演化选择中存活。

牛熊市场状态。在真实世界新闻情景中，我们进一步嵌入两种主要宏观环境：熊市阶段（2022年1月至10月）和牛市阶段（2024年10月至2025年7月）。这些不同的市场状态使我们能够考察演化均衡如何随波动率和情绪状况而变化。如图6所示，稳定的生态配置在不同市场状态之间存在显著差异，表明FinEvo不会坍缩为单一固定结构，而是以符合现实的方式对宏观经济背景作出响应。

总结。 总的来说，受控的人工冲击场景与以实证为基础的真实世界新闻场景构成了一条连贯的验证路径：从机制层面的压力测试到真实市场的稳健性分析。这一设计表明，FinEvo 能够捕捉理论上的演化动态和现实市场中的适应过程，从而全面评估该框架。

G.8 新闻驱动的合成市场的经济动因

本节进一步阐释我们开展新闻驱动实验的经济基础，以及基于大语言模型的智能体在 FinEvo 中所发挥的作用。

不存在偏向大语言模型主导地位的结构性倾向。 FinEvo 并不假定基于大语言模型的智能体具有主导地位，也不会通过结构设计强行确保其占据主导地位；其表现也并非由情绪—收益映射硬编码决定。如图6所示，大语言模型智能体并非始终更优：在牛市状态下，Informer 智能体的表现优于 FinCon 和交易策略；而在熊市状态下，深度 Q 网络取得的收益高于交易策略和思维链。这表明，相对表现取决于不断演化的生态和市场条件，而非对某一特定模型类别的内置偏好。

情绪—收益映射的经济基础。 从新闻情绪到价值/收益冲击的映射，立足于市场微观结构与行为金融学中已确立的结果，而非启发式地选取。秉承Kyle模型中基于信息的交易之精神（Kyle，1985年），信息冲击通过信息型交易员与流动性交易者影响订单流和价格。行为模型表明，情绪与投资者信念偏差可以驱动短周期可预测性与崩盘风险（例如，Barberis等人，1998年）。实证上，重大新闻公告已知会使波动率、跳跃强度和方差动态上涨（Engle与Ng，1993年；Andersen等人，2003年）。在FinEvo中，新闻冲击通过改变智能体的感知价值过程$V_k(t)$进入，这与这些机制一致：情绪并不直接决定价格，而是扰动感知基本面，从而间接影响订单流与收益。

为何使用合成内生价格动态。在新闻实验中，我们有意使用合成生成的内生价格序列，而非重放固定的历史行情序列。我们的目标是研究异质策略与市场之间的闭环演化反馈；真实历史价格无法体现这一点，因为真实价格路径是在一种

未知且未建模的生态中生成的。在 FinEvo 中，新闻冲击作为外生信息输入，而由此产生的价格动态则完全内生，由所有相互作用的智能体（大语言模型、强化学习、深度学习和基于规则的智能体）共同决定。这种设计使我们能够分析不同策略群体如何响应相同的新闻过程，以及演化选择如何随时间推移重塑生态。

G.9 仿真部分的目的

第3节旨在说明，FinEvo 提供了一个通用的动力学框架，可用于研究由相互作用的异质主体所引发的市场层面的演化现象。
这些模拟有三个目的：(1) 验证所提出的演化随机微分方程与连续双向拍卖相结合，能够产生内生的生态动态（选择、优势循环和相变）；(2) 展示系统面对情绪冲击或流动性扰动等扰动时的整体响应；(3) 为市场设计和政策分析提供可控的实验条件。

H 科学研究与交易模型构建

在本附录中，我们将阐明在 FinEvo 的语境下“市场科学研究”的含义。尽管交易模型构建是仿真的一种可能用途，但 FinEvo 旨在支持更广泛的研究议程，聚焦于市场生态系统涌现出的系统性属性。仅仅通过优化智能体在重放历史环境中的损益，无法回答这些问题。

系统性风险与稳定性。 FinEvo 使研究市场整体稳定性如何随策略构成演化而变化成为可能。例如，新“物种”（如基于大语言模型的智能体）的出现，可能改变生态稳定性、波动率持续性，或拥挤诱发的转变（如类似闪崩的动态）。这些现象源于演化随机微分方程中的内生选择、创新和扰动，无法通过训练单个智能体来复现。

政策与监管分析。 监管机构关注的是整体市场的行为，而非个别交易者的利润。FinEvo 提供了一个受控的虚拟环境，用于评估交易税、流动性约束或规则变更等策略如何影响创新率、物种存续和生态多样性。因此，FinEvo 可作为一个政策实验室。

市场设计。 FinEvo 还可用于分析拍卖形式、撮合规则或延迟结构如何影响异质智能体（基于规则、强化学习、大语言模型）之间的生态系统层面共存。这些是关于市场如何作为适应性博弈运行的结构性问题，而非如何构建盈利策略的问题。

摘要。 因此，FinEvo 的主要研究用途是研究市场生态系统的适应性、涌现性和系统性动态——这是市场微观结构和金融经济学研究长期以来的目标。

I 未来工作：提升实证真实性

本附录概述了若干可自然拓展的方向，以提升 FinEvo 的实证真实性。我们目前的贡献侧重于构建演化形式体系——FinEvo 随机微分方程及其选择—创新—扰动动态——并配以具备微观结构感知能力的连续双向拍卖引擎。该框架经过刻意设计，可支撑下文所述的真实性增强。

L2级微观结构校准。 一个重要方向是引入实证限价订单簿统计数据（如 LOBSTER），以校准深度剖面、价差动态、订单到达过程和排队效应。FinEvo 将 CDA机制与演化动态模块化分离，因此可以轻松实现这种整合。

混合回放 + 内生演化。 我们计划引入一种混合市场状态：外部流动性依据历史限价订单簿数据进行回放，而策略群体则在 FinEvoSDE 下内生演化。这既能忠实反映实证数据，也便于进行演化分析。

更丰富的经济反馈机制。通过引入库存约束、市场冲击反馈、交易税、限流规则和自适应市场设计模块，可以进一步提升真实感。这些组件使 FinEvo 不仅可用作演化模型，也可作为评估政策干预和系统性风险的平台。

面向智能体研究的高保真交互环境。更真实的市场生态有助于开展关于强化学习智能体、大语言模型智能体和混合系统的研究。增强后的 FinEvo 环境可提供非平稳的竞争压力、真实的订单簿交互以及无法通过孤立回测获得的涌现集体行为。

科学应用。提升真实感有助于深入研究系统脆弱性、冲击传播、策略集中度以及监管效应对长期生态均衡的影响。这使 FinEvo 与更广泛的市场微观结构和金融经济学研究议程相契合。

开放式策略演化。另一项扩展是允许策略本身随时间演化，而不是固定策略集合。这可以通过遗传算法、神经进化或策略混合机制实现，它们通过变异或重组现有策略来创造新策略。尽管这种开放式演化大幅增强了生态系统的表达能力，但也会导致策略空间迅速扩张，从而增加理论分析的难度。本文着重建立生态博弈形式体系及其种群层面动力学，但 FinEvo 框架完全兼容在未来研究中纳入策略层面的演化。

基于网络的传播与局部扩散。 FinEvo 通过创新项和扰动项对群体层面的扩散进行建模，这与演化博弈论的平均场表述一致；但一个重要的未来方向是纳入基于局部交互的微观层面传播机制。这类动力学包括策略在智能体网络相邻节点间的扩散、图上的局部模仿，以及邻域结构驱动的传染。引入基于网络的传播，将使 FinEvo 能够捕捉空间异质性的扩散模式、集群式采用和局部级联，从而补充当前的平均场演化动力学。FinEvo 随机微分方程采用模块化设计，因此可以轻松通过基于图的转移算子或网络加权创新核扩展演化机制，从而在未来研究中构建更丰富的行为传染和微观结构传播模型。

J 局限性

尽管 FinEvo 提供了一个结合内生市场动力学的统一生态博弈形式体系，但若干建模假设也带来了局限性，同时为未来的改进指明了方向。

演化机制的范围。 演化动力学遵循演化博弈论的平均场传统，在种群层面进行建模。这种方法具有分析上的清晰性，但忽略了基于网络的局部传播。如附录 I 所述，该框架可以自然地扩展为基于图的转移算子，以模拟通过相邻节点进行的扩散。

基于大语言模型的智能体的外部知识。基于大语言模型的智能体可能在预训练期间接触过对某些历史事件的文字描述。由于 FinEvo 生成内生市场动力学，且大语言模型无法访问未来价格或内部微观结构，这种重叠不太可能对我们报告的比较生态结果造成实质性扭曲。

大型语言模型的使用

大语言模型（LLM），例如 ChatGPT，仅用于进行细微的语言润色。所有概念开发、技术贡献、实验和分析均由作者独立完成。