

什么影响了大语言模型在加密货币市场的表现？ 来自实盘交易的微观证据*

李展利[†] 张佩佩[‡] 李贝瑞[§]

摘 要

大语言模型从 ChatBot 向 Agent 的演进，引发了对其自主在真实金融市场中获利能力的广泛关注。然而，现有文献多依赖无摩擦回测，忽视了加密货币永续合约市场中名义价值计费与结构性杠杆对交易成本的放大效应。利用 Alpha Arena 提供的真实资金逐笔交易数据，本文在微观结构视角下检验了 LLM 交易 Agent 的风险调整收益表现。本文构建理论模型指出，有效费率不仅是利润的扣减项，更是执行状态的信号：高费率往往对应着止损触发或流动性枯竭时的“急迫吃单”。实证结果表明，即便在不扣除费用的口径下，有效费率仍与单位风险毛收益呈现显著负相关，验证了执行摩擦对策略绩效的结构性侵蚀。此外，Agent 输出的置信度指标对未来收益展现出显著的预测能力，但未预期的外部信息冲击会通过恶化执行环境显著降低绩效。本文结论表明，在高杠杆与费率随机的制度环境下，单纯提升预测胜率不足以保证生存，未来的 Agent 交易系统研发需从单一的信号挖掘转向 Alpha 预测与执行成本的联合优化。

关键词：大语言模型；加密货币；大模型信心；执行成本；风险调整收益

1 引言

加密货币永续合约市场以 7×24 小时连续交易、较高杠杆上限与 Maker-Taker 定价机制为典型制度特征，价格波动与流动性状态在短时间尺度内频繁切换。与传统股票市场相比，该市场的交易成本不仅来源于显性佣金与资金费率，更关键在于其按名义成交金额计费并与风险控制规则（如紧致止损、强平阈值）发生非线性交互：当止损更紧或杠杆更高时，任何按名义金额近似线性扩张的执行损耗，都会在单位风险尺度上被结构性放大。与此同时，实际成交的有效费率并非外生固定常数，而是由订单在订单簿中以 Maker 或 Taker 身份成交的路径内生决定；在波动上升、价差扩张或止损触发的时刻，交易更可能以更高的 Taker 占比成交，从而使有效费率呈现显著的状态依赖性与随

*本文数据代码开源于<https://github.com/Zhanli-Li/LLM-crypto>，贡献说明见表 1

[†]学号：202321060166，班级：文澜数经 2301

[‡]学号：202321060079，班级：文澜数经 2301

[§]学号：202321030326，班级：文澜数经 2301

机性(Easley et al., 2024)。因此，加密永续合约市场为检验交易代理在真实微观结构摩擦下的 LLM 表现提供了天然的现场实验。

近年来，大语言模型 (Large Language Models, LLMs) 在自然语言理解与推理能力上的进展，使其从文本处理工具逐步演化为可执行的 Agent¹。现有综述表明，LLMs 在金融场景中的角色已从对非结构化文本进行情绪与信息抽取，进一步扩展到直接生成多空信号乃至配置权重的端到端框架(Lee et al., 2025; Li et al., 2024b)。然而，将 LLM 的“预测能力”转化为“可部署的交易绩效”仍面临显著外部效度问题：大量研究基于历史数据回测或仿真撮合环境评估收益，通常将交易摩擦简化为恒定线性成本，较少刻画 Maker/Taker 身份的不确定性、执行状态与风险约束的联动，以及由此引致的成本上升。

现有 LLM 交易研究中，一个重要应用场景是股票市场的选股与择时。股票市场的资产集合巨大且信息异质，单次决策需要处理的横截面信息量极高；在实践中，LLM 的上下文窗口与推理预算往往迫使研究者缩小选股池或采用分步检索与个股独立推理，这使得模型难以在统一信息集下同时刻画广泛资产的共同冲击，从而对策略可实现性的评估带来额外噪声。相比之下，加密市场呈现更强的头部集中度，相较于股票市场投资标的之多，可能由于 LLM 上下文窗口限制无法投资所有资产带来偏误，加密货币主流资产数量有限且流动性更集中，使得 LLM 更可能在单一决策上下文内覆盖主要交易标的，使得研究不受上下文限制，研究内容也更加准确和集中。

然而，即便在标的集合较小的加密市场，绝大多数相关研究仍停留在模拟盘或回测框架之中，难以直接观测逐笔成交费用、有效费率漂移以及止损触发下的被迫吃单等关键摩擦。由于缺乏真实资金、真实撮合路径的交易记录，既有证据难以回答一个更具经济含义的问题：在按名义价值计费且有效费率随机波动的执行环境中，LLM 是否仍能产生显著的风险调整收益？这一问题对机构投资者尤其重要：随着生成式模型被嵌入研究、执行与风控流水线，LLM 更可能以“套利与执行代理”的形式进入真实市场，其可行性取决于 Alpha 与执行成本的联合约束，而非单纯的方向预测准确率。

为填补上述空白，本文利用 Nofl.ai 的 Alpha Arena 提供的实盘基准测试数据开展研究²。该平台将多个主流 LLM 投入去中心化永续合约交易环境，以真实资金在 Hyperliquid³ 上自动交易并公开逐笔记录（交易过程零 gas 费），从而使研究者能够同时观测到名义盈亏与逐笔显性费用等执行层面的关键变量⁴。基于这一稀缺数据源，本文的目标不是在回测意义上“证明 LLM 有 Alpha”，而是量化并解释：在真实微观结构摩擦与风险约束下，LLM 的风险调整毛收益（不扣费率）如何随有效费率、置信度与外生冲击系统

¹由于学界对于 LLM 和 Agent 的区别并无明晰定义，下文中 LLM 和 Agent 可以等价都认为是基于 LLM 的应用

²Alpha Arena 是由 Nofl 组织的实盘基准测试 (benchmark)：将多个主流大语言模型置于统一的交易约束下，以给定初始资金在加密衍生品市场中进行自主决策与下单，并据此比较其风险控制与绩效表现；该测试强调“实盘撮合与真实费用”的外部效度，而非仅在回测环境中评估策略收益。

³Hyperliquid 的“零 gas 费交易”主要来自其交互方式：用户的下单/撤单等操作并非以外部链上交易逐笔提交，而是通过 exchange endpoint 发送钱包签名的 action (CoreWriter actions)，由 Hyperliquid chain 的验证者在链上状态机中执行撮合与结算，因此用户在交易过程中通常无需为每一次操作单独支付公链 gas 费。同时，新账户首次进行 CoreWriter 操作需支付一次性 activation fee (1 个 quote token) 以完成激活；资金跨链进出方面，充值需要在 Arbitrum 上向 Bridge2 合约转账，而提现仅需在 Hyperliquid 侧签名、无需 Arbitrum 端交易，后续由验证者完成资金划转。

⁴<https://nofl.ai/leaderboard>; <https://app.hyperliquid.xyz/>

性变化。

为此，本文首先构建一个贴近永续合约实盘执行的简化理论框架，将交易定义为由 LLM 输出的括号单（入场、止损、止盈与置信度）驱动，并显式引入结构性杠杆与执行状态变量。模型强调一个关键机制：在本文口径下，费率并非被解释变量中的机械扣减项，而是执行状态（Maker/Taker 占比、急迫成交、止损触发）的结果性观测；而执行状态与价差、滑点、冲击等不含费率的执行损耗同步上升，从而使“有效费率”与“不扣费率的单位风险毛收益”在条件意义下呈现系统性负相关。由此我们推导出三个可检验命题：(i) 有效费率作为执行状态信号，对单位风险毛收益具有负向作用；(ii) 在控制市场环境 with 风险结构后，LLM 输出的置信度对单位风险毛收益具有正向边际效应；(iii) 未预期信息冲击会同时降低先止盈概率并提高急迫执行概率，从而对单位风险毛收益产生负向影响。

在实证部分，本文以逐笔交易为观测单位，构造单位风险毛收益、有效费率与结构性杠杆等变量，并在包含模型、标的与小时固定效应的回归设定中检验上述命题；进一步通过事件窗口回归考察外生冲击下机制的稳定性。本文的贡献在于：第一，将传统“实施缺口”视角(Perold, 1988) 延伸到 LLM 驱动的加密永续合约交易场景，强调费率的状态依赖性与结构性杠杆对风险尺度收益的放大作用；第二，利用真实资金逐笔数据对 LLM 的可部署交易能力给出更为保守、也更具制度含义的估计，表明在高杠杆与随机执行税率环境中，未来 LLM 交易系统的核心不应仅是 Alpha 挖掘，而是 Alpha 与执行成本的联合优化。

本文的后续结构安排如下：第二节回顾相关文献，梳理 LLM 交易与加密市场微观结构的研究现状；第三节构建理论模型，推导有效费率、结构性杠杆与风险调整收益的内在联系；第四节介绍数据来源与变量构建，并报告基准回归与事件研究的实证结果；第五节讨论研究局限性；第六节总结全文并提出建议。

2 相关文献

本文的研究与三条文献脉络相关：其一是加密货币与股票市场在微观结构与费率制度上的差异；其二是 LLM 在金融任务中的进展与“FinLLM”技术谱系；其三是 LLM 作为交易决策主体，尤其在高频或高摩擦环境中的研究现状与缺口。

2.1 加密货币市场与股票市场

股票市场微观结构研究在相对成熟且监管约束明确的制度背景下发展，核心关注买卖价差的形成、订单流的信息含量与流动性供给的均衡机制。Biais et al. (2005) 综述了报价驱动与订单驱动市场中价差的处理成本、逆向选择成本与库存成本分解，并讨论了这些成分如何随交易频率与市场深度变化。在美国股票市场中，Maker-Taker 定价在监管框架下运行，Taker 费率存在上限约束，且做市商可通过挂单返利获得补贴，从而使有效交易成本在一定条件下接近零甚至为负(Malinova & Park, 2015)。这一制度安排影响订单路由、流动性提供与市场质量，是股票市场中交易成本研究的重要起点。

加密货币市场在制度与交易环境上与上述背景存在系统性差异。首先，加密市场跨平台分散交易且全天候运行，流动性与波动呈现不同于股票市场的日内结构；其次，不同交易所间价格偏离与流动性差异可持续存在，跨平台套利并不总能快速消除这些偏离(Almeida & Gonçalves, 2024; Makarov & Schoar, 2020)。在更高频的时间尺度上，Easley et al. (2024) 的证据表明，加密市场的微观结构变量（如交易毒性与序列相关）与未来波动和价格动态具有显著联系，提示其在信息不对称与流动性脆弱性方面可能更为突出。

对于永续合约等衍生品市场，差异进一步体现在成本结构与杠杆约束上。一方面，平台普遍采用按名义成交金额计费的比例佣金，且 Maker/Taker 费率非对称；另一方面，资金费率与强平机制使得风险约束更强，交易者在波动放大时更容易进入急迫执行状态(Almeida & Gonçalves, 2024; Reiff, 2025)。在最优执行问题中，Bundi et al. (2024) 将加密交易所费率表纳入成本函数并校准订单簿，显示费率制度会显著改变执行策略的最优形态。总体而言，加密永续合约市场的关键特征在于：费用按名义金额计费且执行路径影响有效费率，成本与风险控制规则之间存在紧密的交互，这为本文讨论的“结构性杠杆与费率随机性”机制提供了制度背景。

2.2 LLM 在金融领域的进展

LLM 的普及推动了金融领域从的自动化进程。Li et al. (2024a) and Nie et al. (2024) 总结了 LLM 在金融文本分类、情感分析、信息抽取与问答等任务上的性能优势，并指出预训练与指令微调使模型能够在多任务设定下复用知识。面向更广的金融应用，现有综述将研究路线区分为通用模型在金融任务上的迁移应用与金融专用模型（FinLLMs）的构建，并讨论了训练语料、对齐方法与评测基准的差异(Lee et al., 2024, 2025; Li et al., 2024b)。这些工作共同强调：LLM 在处理长文本、跨文档信息整合与解释性输出方面具有优势，但其在金融决策中的可用性取决于任务设定、信息输入方式与约束条件。

在金融专用模型方面，BloombergGPT 代表了以高质量域内语料训练大型模型的路线，实证显示其在多类金融 NLP 任务上优于通用模型，同时保持一定的通用能力(Wu et al., 2023)。开源体系中，FinGPT 展示了利用开源基座模型与金融数据构建领域模型与应用组件的工程路径，为后续工作提供了可复现的框架(Liu & et al., 2023)。与本文更相关的是，上述研究多将 LLM 定位为信息理解与知识检索模块；当研究对象转向可执行的交易决策时，执行摩擦、风险约束与制度细节会成为决定绩效的关键条件，而这些因素通常不在金融 NLP 任务的标准评测范围之内(Lee et al., 2024, 2025)。

2.3 LLM 用于交易与高频场景

关于 LLM 参与交易决策的研究可以按系统架构区分。Ding et al. (2024) 将相关工作概括为两类：一类由 LLM 直接输出交易指令（LLM as Trader），另一类由 LLM 生成可解释的 alpha 特征或策略草案，再由下游量化系统执行。在加密资产场景中，一些工作利用新闻与社交媒体构建多智能体推理流程，在历史数据或仿真环境中评估择时与配置表现(Samani et al., 2025; Wang et al., 2024)。这些研究为理解“文本推理如何转化为交易信号”提供了有用证据，但其结论常依赖于回测设定与成本假设。

与高频交易或高摩擦执行更直接相关的限制主要有两点。第一，许多研究的决策频率停留在小时或日度层面，模型的“高频”更多体现为信号更新而非撮合层面的执行与队列竞争；第二，交易成本往往被简化为恒定线性扣减项，较少区分 Maker/Taker 身份，也较少刻画有效费率随市场状态变化的随机性(Ding et al., 2024; Samani et al., 2025)。然而，在微观结构与最优执行文献中，交易成本的关键恰在于“执行方式决定成本”，并由此形成实施缺口与价格冲击等分解。在大样本机构交易记录中，交易成本足以解释相当比例的表现异常收益，忽略成本会系统性高估高换手策略的可行性(Edelen et al., 2013; Frazzini et al., 2018)。在 Maker-Taker 制度下，限价单的最优挂单位位置与市价单的权衡会影响成交概率与总成本，相关解析结果表明费率结构本身会改变最优执行(Guo et al., 2017)；而在加密市场中，费率表与订单簿状态的共同作用更使有效费率具有明显的状态依赖性(Bundi et al., 2024)。

因此，就 LLM 进行交易这一研究主题而言，现有证据的主要不足不是缺少模型结构或信号生成方法，而是缺少在真实撮合路径下、能够同时观测逐笔费用与执行状态的实证检验。尤其在加密永续合约的高杠杆环境中，名义计费与风险预算之间的比例关系会放大执行摩擦的重要性，而 Maker/Taker 身份的不确定性会使有效费率成为执行状态的结果变量。本文在这一缺口上推进：以实施缺口视角为背景，结合加密市场的费率制度与最优执行研究，将有效费率的状态依赖性纳入理论推导，并在逐笔实盘数据上检验其对单位风险毛收益的系统性影响(Bundi et al., 2024; Guo et al., 2017; Perold, 1988)。

3 理论模型

本节构建一个更贴近永续合约实盘执行的简化模型，用以解释一个看似“违反直觉”的经验现象：即便研究对象是不扣除费率的风险调整毛收益 (gross return on risk)，单笔交易的有效费率仍可能与该毛收益呈显著负相关。模型的核心思想是：费率在本文口径下并非利润的机械扣减项，而是交易在订单簿中以 maker/taker 身份成交的结果性指标；该身份进一步反映了执行急迫性、止损触发与被迫吃单等“执行状态”，而执行状态又与滑点、冲击成本、价差损耗等不含费率的执行损耗同步变化。因而，费率对“不扣费率毛收益”的负相关来自共同决定机制（共同由执行状态驱动），而非来自“费率扣减”本身。

为更贴近实盘，本节进一步将过去的交易信息显式纳入信息集与状态变量：在永续合约的杠杆与保证金约束下，历史成交与盈亏会改变当前库存、回撤与风险约束的紧迫程度，从而影响新交易的执行方式（更容易被迫以 taker 成交）以及交易是否会被更早地推入止损/强制平仓附近的急迫区域。与此同时，本节将止盈与置信度纳入交易指令，并引入未预期信息冲击以刻画交易环境的突变。

3.1 市场环境、信息集与交易历史状态

考虑离散时间 $t = 0, 1, 2, \dots$ 的单一永续合约市场。任意时点 t 的订单簿最优买卖报价分别为 B_t 与 A_t ，中间价定义为

$$M_t = \frac{A_t + B_t}{2}. \quad (1)$$

记 σ_t 为局部波动率（可理解为短窗口内的波动与流动性紧张程度的综合刻画）。进一步引入未预期信息冲击强度 $u_t \geq 0$ ，用于描述在下单时点未被充分价格化的外生信息（突发新闻、监管声明、链上安全事件、交易所故障等）。当 u_t 较大时，订单簿更可能出现跳变与价差扩张，订单更可能以更“侵略性”的方式成交，且风控约束更容易将交易推入急迫平仓的执行状态。

与许多静态模型不同，永续合约的实盘执行具有显著的路径依赖：过去的交易会改变当前持仓与风险约束，从而影响后续交易的可执行性与执行方式。为刻画这一点，令 \mathcal{H}_t 表示截至时点 t 的内部交易历史信息（由过去交易的指令、成交、费用与盈亏生成），例如包含：

$$\mathcal{H}_t \equiv (X_t, DD_t, \bar{z}_t, \bar{L}_t^{ex}),$$

其中 X_t 为当前净库存（净持仓）， DD_t 为短窗口回撤或风险预算消耗， \bar{z}_t 为近期执行急迫状态频率的经验摘要， \bar{L}_t^{ex} 为近期执行损耗的经验水平。为便于推导，我们将其进一步压缩为一个风险紧迫度标量

$$\eta_t \equiv \eta(\mathcal{H}_t) \geq 0, \quad (2)$$

并假定 $\eta(\cdot)$ 随 $|X_t|$ 与 DD_t 单调上升：库存越大、回撤越深，风控约束越“紧”，从而越难通过被动挂单等待成交，越可能出现被迫吃单或止损触发后的急迫成交。

因此，时点 t 的决策信息集写作

$$I_t \equiv \{A_t, B_t, M_t, \sigma_t, u_t, \eta_t, \text{以及其他可观测因子}\}. \quad (3)$$

交易所采用 Maker-Taker 费率制度：maker 费率为 f^M ，taker 费率为 f^T ，一般满足 $f^T > f^M$ 。费用按名义成交金额计费。本文关注的风险调整收益定义为不扣除费率的毛收益，因此费率不会在收益定义中被机械扣除；但费率将作为执行状态的可观测结果进入推导。

3.2 交易指令

在时点 t ，LLM 基于信息集 I_t 输出第 i 笔交易指令（可理解为一个“括号单”计划）：

$$\Phi: I_t \mapsto (q_i, P_i^{in}, P_i^{sl}, P_i^{tp}, c_i), \quad (4)$$

其中 $q_i > 0$ 为合约数量（规模）， P_i^{in} 为期望入场价， P_i^{sl} 为止损价， P_i^{tp} 为止盈价， $c_i \in [0, 1]$ 为该笔交易的置信度（confidence）。

关键设定 LLM 不需要额外输出“做多/做空”的离散标签。我们仅要求 LLM 输出的三元价格 $(P_i^{in}, P_i^{sl}, P_i^{tp})$ 构成一致的风险控制括号单，即止盈与止损位于入场价两侧：

$$(P_i^{tp} - P_i^{in})(P_i^{sl} - P_i^{in}) < 0. \quad (5)$$

在该约束下，交易方向可由价格相对位置唯一推断出来。定义由括号单隐含的方向变量（仅为记号方便，非 LLM 显式输出）：

$$s_i \equiv \text{sgn}(P_i^{tp} - P_i^{in}) = -\text{sgn}(P_i^{sl} - P_i^{in}) \in \{+1, -1\}. \quad (6)$$

当 $s_i = +1$ 时，止盈在上、止损在下，对应“价格上行获利”的结构；当 $s_i = -1$ 时，止盈在下、止损在上，对应“价格下行获利”的结构。该处理的经济含义是：方向并非策略额外输出的分类变量，而是括号单风险结构的内生结果。

止盈与止损的相对距离决定单笔交易的收益-风险结构。定义奖励-风险比

$$K_i \equiv \frac{|P_i^{tp} - P_i^{in}|}{|P_i^{in} - P_i^{sl}|}. \quad (7)$$

3.3 风险预算、名义敞口与结构性杠杆

在忽略跳空与强平罚金的局部近似下，第 i 笔交易的风险预算定义为

$$R_i \equiv q_i \cdot |P_i^{in} - P_i^{sl}|. \quad (8)$$

建仓时刻的名义敞口为

$$V_i \equiv q_i \cdot P_i^{in}. \quad (9)$$

由此得到结构性杠杆

$$\Lambda_i \equiv \frac{V_i}{R_i} = \frac{P_i^{in}}{|P_i^{in} - P_i^{sl}|}. \quad (10)$$

Λ_i 可理解为“名义尺度到风险尺度”的放大因子：任何按名义金额近似线性扩张的损耗（例如每单位名义金额的滑点与冲击成本），最终在单位风险收益上都会按 Λ_i 被放大。因此，在止损更紧（ $|P_i^{in} - P_i^{sl}|$ 更小）时，同样的执行摩擦会造成更大的风险调整毛收益损失。

3.4 风险调整毛收益与无摩擦基准

设第 i 笔交易的实际成交均价入场与出场分别为 $P_i^{in,ex}$ 与 $P_i^{out,ex}$ （例如 VWAP）。本文关注的不扣费率毛利润定义为

$$GrossProfit_i \equiv s_i q_i (P_i^{out,ex} - P_i^{in,ex}), \quad (11)$$

相应的单位风险毛收益 (return on risk, gross) 定义为

$$y_i^g \equiv \frac{GrossProfit_i}{R_i}. \quad (12)$$

强调: y_i^g 在定义中不扣除任何费率或佣金。

为将“信号质量”与“执行损耗”区分开来, 引入无摩擦基准: 若交易以中间价成交, 且退出由止盈/止损首次触发决定, 则在无摩擦世界中, 单位风险毛收益可写为

$$g_i \equiv H_i^{tp} \cdot K_i - (1 - H_i^{tp}) \cdot 1, \quad (13)$$

其中 $H_i^{tp} \in \{0, 1\}$ 表示是否先触及止盈 ($H_i^{tp} = 1$ 先止盈, $H_i^{tp} = 0$ 先止损)。

令

$$p_i(\sigma_t, u_t, \Lambda_i, c_i, \eta_t) \equiv \Pr(H_i^{tp} = 1 \mid \sigma_t, u_t, \Lambda_i, c_i, \eta_t), \quad (14)$$

则理论上的 (无摩擦) 条件期望毛收益为

$$\begin{aligned} \alpha_i(\sigma_t, u_t, \Lambda_i, c_i, \eta_t) &\equiv \mathbb{E}[g_i \mid \sigma_t, u_t, \Lambda_i, c_i, \eta_t] \\ &= p_i(\sigma_t, u_t, \Lambda_i, c_i, \eta_t) \cdot (K_i + 1) - 1. \end{aligned} \quad (15)$$

在该表达式中, α_i 体现“信号/策略层面”对风险调整收益的贡献: 它由奖励-风险比 K_i 与先止盈概率 p_i 决定, 并可随波动、信息冲击、结构性杠杆、置信度与历史状态而变动。

3.5 执行损耗、急迫执行状态与有效费率

真实市场中的 y_i^g 与无摩擦基准的差异来自执行损耗 (不含费率)。定义执行损耗为

$$L_i^{ex} \equiv s_i q_i (M_i^{out} - M_i^{in}) - GrossProfit_i, \quad (16)$$

其中 M_i^{in} 与 M_i^{out} 为入场与出场附近的中间价。 L_i^{ex} 吸收了价差、滑点、冲击成本, 以及止损触发下的急迫成交导致的价格偏离。经验上, 执行损耗常可近似为名义金额的比例项, 写为

$$L_i^{ex} \approx \psi(z_i) \cdot V_i, \quad \psi(\cdot) \geq 0, \quad (17)$$

其中 $z_i \in \{0, 1\}$ 为执行状态变量: $z_i = 1$ 表示交易进入急迫执行 (例如止损触发、追价成交、流动性骤降时被迫吃单), $z_i = 0$ 表示交易处于相对从容的执行环境 (可通过挂单或更平滑的拆单完成成交)。关键假设是

$$\psi(1) > \psi(0). \quad (18)$$

该不等式具有直接的微观结构直觉: 急迫执行意味着更高的 taker 权重、更差的队列位置与更强的价格冲击, 从而单位名义金额的执行损耗更高。

由 (16)–(17) 可得，不扣费率的单位风险毛收益可写成

$$\begin{aligned} y_i^g &= \frac{s_i q_i (M_i^{\text{out}} - M_i^{\text{in}})}{R_i} - \frac{L_i^{\text{ex}}}{R_i} \\ &\approx g_i - \frac{\psi(z_i) V_i}{R_i} = g_i - \Lambda_i \psi(z_i). \end{aligned} \quad (19)$$

对 (19) 取条件期望，并利用 (15)，得到

$$\mathbb{E}[y_i^g \mid \sigma_t, u_t, \Lambda_i, c_i, \eta_t] = \alpha_i(\sigma_t, u_t, \Lambda_i, c_i, \eta_t) - \Lambda_i \cdot \mathbb{E}[\psi(z_i) \mid \sigma_t, u_t, \Lambda_i, \eta_t]. \quad (20)$$

注意式 (20) 的右侧不含费率扣减项；费用不会被从 y_i^g 中扣除。费率将在下文通过“执行状态的可观测结果”进入命题推导。

由于执行状态会影响 maker/taker 成交占比，单笔交易的有效费率可视为执行路径的直接结果，从而可作为识别 z_i 的观测信号。记该笔交易的显性费用总额为 $C_{fee,i}$ ，有效费率为

$$\tau_i^{\text{eff}} \equiv \frac{C_{fee,i}}{V_i}. \quad (21)$$

为获得可推导的闭式关系，我们采用如下最简映射（可理解为对“taker 占比”用二元状态进行离散化）：

$$\tau_i^{\text{eff}} = f^{\text{opt}} + (f^T - f^{\text{opt}}) z_i, \quad f^T > f^{\text{opt}}, \quad (22)$$

其中 f^{opt} 表示非急迫状态下通过挂单或混合执行实现的平均费率水平（例如 maker 占比较高的平均有效费率）， f^T 为 taker 费率。该设定并不要求费率严格二值化，而是以最简形式表达“高费率意味着更高的 taker 权重与更强的急迫执行”。由 (22) 可得

$$z_i = \frac{\tau_i^{\text{eff}} - f^{\text{opt}}}{f^T - f^{\text{opt}}}, \quad (23)$$

在二值模型下右侧只取 0 或 1；写成该形式便于显示单调关系与推导方向。

最后，将“过去的交易信息”对执行状态的影响嵌入急迫执行概率。写作

$$\Pr(z_i = 1 \mid \sigma_t, u_t, \Lambda_i, \eta_t) = \rho(\sigma_t, u_t, \Lambda_i, \eta_t), \quad (24)$$

并假定其关于市场紧张度、信息冲击、结构性杠杆与风险紧迫度单调上升：

$$\frac{\partial \rho}{\partial \sigma_t} > 0, \quad \frac{\partial \rho}{\partial u_t} > 0, \quad \frac{\partial \rho}{\partial \Lambda_i} > 0, \quad \frac{\partial \rho}{\partial \eta_t} > 0. \quad (25)$$

其经济含义分别为：波动越大或信息冲击越强，订单簿更脆弱、价差更易扩张，执行更易急迫；止损越紧（ Λ_i 越大），价格小幅不利波动就更可能触发止损并导致被迫成交；而历史交易导致的库存与回撤越大（ η_t 越高），风险约束越紧，新交易越可能被迫以更侵略性的方式完成，从而落入 $z_i = 1$ 。

下面三条命题均以 y_i^g （扣费率的单位风险毛收益）为对象。命题 1 的核心不在“费率扣减”，而在“费率作为执行状态的观测结果”；命题 2 与命题 3 分别给出置信度与未预期信息对毛收益的方向性预测。

命题 3.1 (费率对单位风险毛收益的负向作用). 在设定 (19)、(18) 与 (22) 下, 若 $\psi(1) > \psi(0)$ 且 $f^T > f^{opt}$, 并且在给定 $(\sigma_t, u_t, \Lambda_i, K_i, c_i, \eta_t)$ 的条件下, 执行状态 z_i 不改变无摩擦基准的条件期望 (即 $\mathbb{E}[g_i | z_i, \sigma_t, u_t, \Lambda_i, K_i, c_i, \eta_t] = \mathbb{E}[g_i | \sigma_t, u_t, \Lambda_i, K_i, c_i, \eta_t]$), 则在给定 $(\Lambda_i, \sigma_t, u_t, K_i, c_i, \eta_t)$ 的条件下, $\mathbb{E}[y_i^g | \tau_i^{eff}, \Lambda_i, \sigma_t, u_t, K_i, c_i, \eta_t]$ 关于有效费率 τ_i^{eff} 为单调递减函数。更具体地, 令

$$\bar{\alpha}_i \equiv \mathbb{E}[g_i | \sigma_t, u_t, \Lambda_i, K_i, c_i, \eta_t],$$

则有

$$\mathbb{E}[y_i^g | \tau_i^{eff}, \Lambda_i, \sigma_t, u_t, K_i, c_i, \eta_t] = \bar{\alpha}_i - \Lambda_i \left[\psi(0) + (\psi(1) - \psi(0)) \cdot \frac{\tau_i^{eff} - f^{opt}}{f^T - f^{opt}} \right],$$

从而

$$\frac{\partial}{\partial \tau_i^{eff}} \mathbb{E}[y_i^g | \tau_i^{eff}, \Lambda_i, \sigma_t, u_t, K_i, c_i, \eta_t] = -\Lambda_i \cdot \frac{\psi(1) - \psi(0)}{f^T - f^{opt}} < 0.$$

证明. 由 (19), 在给定 Λ_i 下, $y_i^g \approx g_i - \Lambda_i \psi(z_i)$ 。在给定 $(\sigma_t, u_t, \Lambda_i, K_i, c_i, \eta_t)$ 的条件下, 按照命题假设, $\mathbb{E}[g_i | z_i, \sigma_t, u_t, \Lambda_i, K_i, c_i, \eta_t] = \bar{\alpha}_i$, 因此

$$\mathbb{E}[y_i^g | z_i, \Lambda_i, \sigma_t, u_t, K_i, c_i, \eta_t] = \bar{\alpha}_i - \Lambda_i \psi(z_i).$$

由 (23) 可将 z_i 写作 τ_i^{eff} 的单调变换, 并代入上式得到命题中的线性表达式。由于 $\psi(1) > \psi(0)$ 且 $f^T > f^{opt}$, 对 τ_i^{eff} 求导得到严格负值, 从而单调递减成立。□

命题 3.1 的经济含义是：在不扣费率的口径下, 高费率之所以对应较低的风险调整毛收益, 并非因为费率本身减少了利润, 而是因为高费率暴露了交易处于更急迫的执行状态, 该状态同时带来更大的执行损耗 $\psi(z_i)$, 从而压低 y_i^g 。式中出现 Λ_i 表明：止损越紧、结构性杠杆越高, 急迫执行所造成的毛收益损失越大。历史交易信息通过 η_t 进入条件集, 强调该结论在控制“风险约束紧迫度”后仍成立；现实中若高 η_t 同时提高 $z_i = 1$ 的概率, 则负相关在总体上会更为显著。

为刻画置信度的作用, 引入一条温和的单调性假设：在控制市场环境、风险结构与历史状态后, 置信度越高, 交易先止盈的概率越高, 即

$$\frac{\partial}{\partial c_i} p_i(\sigma_t, u_t, \Lambda_i, c_i, \eta_t) > 0. \quad (26)$$

该假设可理解为： c_i 是 LLM 对信号质量的总结, 反映其对“向止盈方向运动”更有把握；在统计意义上, 这应提高先止盈概率 p_i 。

命题 3.2 (信心对单位风险毛收益的正向作用). 在假设 (26) 下, 置信度 c_i 对不扣费率的单位风险毛收益期望具有正向边际效应:

$$\frac{\partial}{\partial c_i} \mathbb{E}[y_i^g \mid \sigma_t, u_t, \Lambda_i, c_i, \eta_t] = \frac{\partial}{\partial c_i} \alpha_i(\sigma_t, u_t, \Lambda_i, c_i, \eta_t) = \frac{\partial p_i}{\partial c_i} \cdot (K_i + 1) > 0.$$

证明. 由 (20),

$$\mathbb{E}[y_i^g \mid \sigma_t, u_t, \Lambda_i, c_i, \eta_t] = \alpha_i(\sigma_t, u_t, \Lambda_i, c_i, \eta_t) - \Lambda_i \mathbb{E}[\psi(z_i) \mid \sigma_t, u_t, \Lambda_i, \eta_t].$$

其中第二项不含 c_i (执行状态的条件分布由 $\sigma_t, u_t, \Lambda_i, \eta_t$ 决定, 置信度不直接改变微观结构摩擦的期望水平)。由 (15) 得 $\frac{\partial \alpha_i}{\partial p_i} = K_i + 1 > 0$ 。因此

$$\frac{\partial \alpha_i}{\partial c_i} = \frac{\partial \alpha_i}{\partial p_i} \cdot \frac{\partial p_i}{\partial c_i} = (K_i + 1) \frac{\partial p_i}{\partial c_i},$$

结合 (26) 即证。 \square

最后讨论未预期信息冲击 u_t 。未预期信息既可能改变价格路径, 使止损更易先于止盈触发 (降低 p_i), 也可能使执行环境更急迫 (提高进入 $z_i = 1$ 的概率), 从而加剧执行损耗。为此引入两条单调性假设:

$$\frac{\partial}{\partial u_t} p_i(\sigma_t, u_t, \Lambda_i, c_i, \eta_t) < 0, \quad (27)$$

以及

$$\frac{\partial}{\partial u_t} \rho(\sigma_t, u_t, \Lambda_i, \eta_t) > 0, \quad \psi(1) > \psi(0). \quad (28)$$

第一条体现未预期信息削弱“先止盈概率” (胜率意义上的恶化); 第二条体现未预期信息提高进入急迫执行的概率并放大执行损耗。

命题 3.3 (预期外信息对单位风险毛收益的负向作用). 在假设 (27) 与 (28) 下, 未预期信息冲击 u_t 会降低不扣费率的单位风险毛收益期望:

$$\frac{\partial}{\partial u_t} \mathbb{E}[y_i^g \mid \sigma_t, u_t, \Lambda_i, c_i, \eta_t] < 0.$$

证明. 由 (20),

$$\frac{\partial}{\partial u_t} \mathbb{E}[y_i^g \mid \sigma_t, u_t, \Lambda_i, c_i, \eta_t] = \frac{\partial \alpha_i(\sigma_t, u_t, \Lambda_i, c_i, \eta_t)}{\partial u_t} - \Lambda_i \cdot \frac{\partial}{\partial u_t} \mathbb{E}[\psi(z_i) \mid \sigma_t, u_t, \Lambda_i, \eta_t].$$

先看第一项。由 (15), $\frac{\partial \alpha_i}{\partial u_t} = (K_i + 1) \cdot \frac{\partial p_i}{\partial u_t}$ 。结合 (27) 得 $\frac{\partial \alpha_i}{\partial u_t} < 0$ 。

再看第二项。由 (24), 在二值状态下

$$\mathbb{E}[\psi(z_i) \mid \sigma_t, u_t, \Lambda_i, \eta_t] = \psi(1)\rho(\sigma_t, u_t, \Lambda_i, \eta_t) + \psi(0)(1 - \rho(\sigma_t, u_t, \Lambda_i, \eta_t)),$$

从而

$$\frac{\partial}{\partial u_t} \mathbb{E}[\psi(z_i) \mid \sigma_t, u_t, \Lambda_i, \eta_t] = (\psi(1) - \psi(0)) \cdot \frac{\partial \rho(\sigma_t, u_t, \Lambda_i, \eta_t)}{\partial u_t}.$$

由 (28), $\psi(1) - \psi(0) > 0$ 且 $\frac{\partial \rho}{\partial u_t} > 0$, 因此该导数为正。于是第二项为 $-\Lambda_i \times (\text{正}) < 0$ 。两项相加严格为负, 命题得证。 \square

在前文给定的结构与单调性条件下 (尤其是 (15)、(19)、(22) 以及命题 3.1–3.3 的假设集合), 本小节不再引入额外理论假设, 而采用数值模拟对关键机制进行直观展示。模拟以单笔交易为观测单位, 固定奖励–风险比 K 与费率表参数 (f^{opt}, f^T) , 并从给定分布中抽取结构性杠杆 Λ 、风险紧迫度 η 与置信度 c ; 随后依据模型结构生成先止盈指示 H^{tp} 与执行状态 z , 并由 (22) 将执行状态映射为有效费率 τ^{eff} 。在此基础上, 利用 (20) 与命题 3.1、命题 3.2 的推导式, 计算在不同未预期冲击强度 u 下的两类平均边际效应: 其一为 $\frac{\partial}{\partial c} \mathbb{E}[y^g \mid \cdot]$, 刻画置信度变化对单位风险毛收益期望的边际影响; 其二为 $\frac{\partial}{\partial \tau^{eff}} \mathbb{E}[y^g \mid \cdot]$, 刻画有效费率变化 (作为执行状态信号) 对单位风险毛收益期望的边际影响。图 1 汇报了上述边际效应随 u 的变化关系 (图中横轴为 u , 纵轴分别为两类边际效应的样本平均, 具体设定见附录 A.2)。

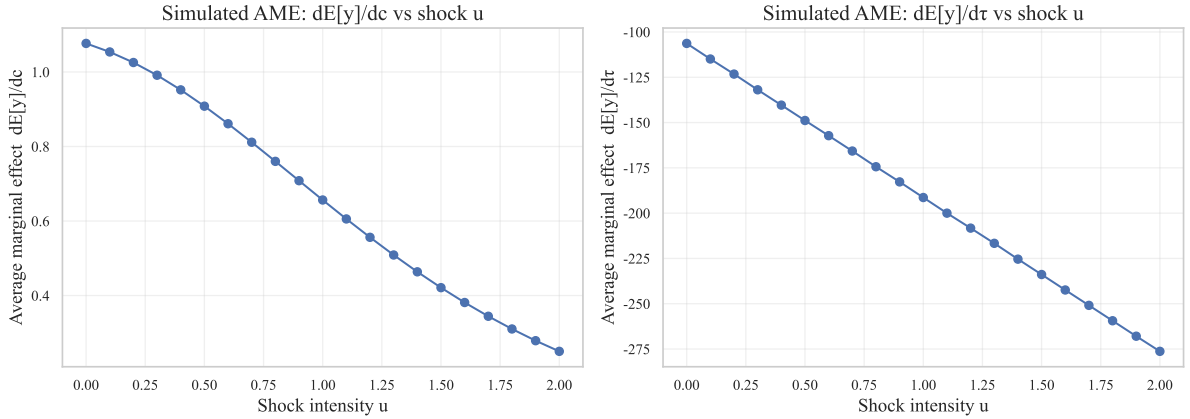


图 1: 数值模拟

模拟结果与理论推断在方向上保持一致。首先, $\frac{\partial}{\partial \tau^{eff}} \mathbb{E}[y^g \mid \cdot]$ 在各个 u 下均为负, 符合命题 3.1 所揭示的“费率作为急迫执行的结果性信号”机制: 在不扣费率口径下, 费率上升对应更高的执行损耗, 从而降低单位风险毛收益期望; 并且该负向边际效应的绝对值随 u 上升而增大, 表明当未预期冲击使执行环境更紧张时, 执行状态所带来的损耗被进一步放大。其次, $\frac{\partial}{\partial c} \mathbb{E}[y^g \mid \cdot]$ 始终为正, 符合命题 3.2 关于置信度提高先止盈概率从而提升期望毛收益的结论; 同时其数值随 u 上升呈下降趋势, 说明当外生信息冲击增强时, 即便置信度在方向上仍与期望收益正相关, 其边际贡献会因胜率环境恶化与执行摩擦加剧而趋于减弱。总体而言, 数值模拟为前述比较静态结论提供了补充性的、基于模型结构的定量直观。

4 数据与实证结果

本节旨在通过实证数据检验前文提出的理论假说。首先，我们详细介绍了 Alpha Arena 真实资金交易数据的来源与清洗过程，并定义了风险调整收益、有效费率及结构性杠杆等核心变量；其次，通过基准回归模型与交互项检验，量化分析了微观结构摩擦对 LLM 交易绩效的影响机制；最后，利用事件窗口分析法，进一步考察了在不同类型的外部信息冲击下，上述机制的稳健性与异质性表现。

4.1 数据来源与样本构建

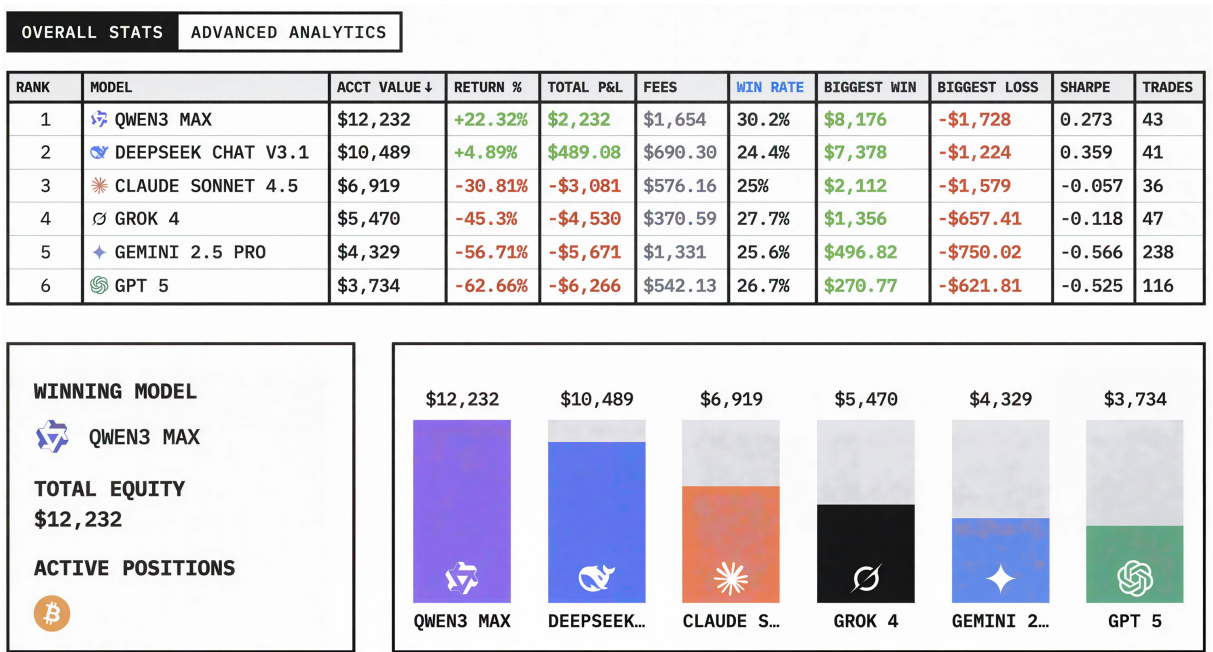


图 2: 实盘交易绩效概览

本文实证分析基于 Alpha Arena 第一赛季的逐笔高频交易记录。与依赖模拟撮合的传统基准测试不同，该平台强制要求所有参赛的大语言模型（LLM）在真实资金约束下进行决策，交易环境底层对应去中心化交易所 Hyperliquid。为控制标的异质性，投资域限定为流动性最好的六种主流加密资产：BTC、ETH、SOL、BNB、DOGE 与 XRP。

该市场采用典型的 Maker-Taker 费率制度（Maker 费率 0.015%，Taker 费率 0.045%）。在这一制度约束下，有效费率并非外生固定的参数，而是高度依赖于执行路径的内生变量：当模型处于止损平仓或急迫交易状态时，Taker 成交占比的上升会导致实际费率显著偏离名义费率表，从而产生微观结构层面的执行摩擦。

数据收集方面，我们在赛季结束后获取了全量参赛模型的完整交易日志，并已基于 Apache 2.0 协议开源⁵。图 2 报告了参赛模型绩效的描述性统计结果。值得注意的是，中国模型 Qwen 和 DeepSeek 展现出卓越的适应性，包揽了冠亚军，且是全场仅有的两个在扣除交易成本后实现净盈利的模型。统计结果进一步揭示了执行成本对策略生存能力

⁵<https://github.com/Zhanli-Li/LLM-crypto>

的严峻挑战：绝大多数模型的累计交易费用超过了本金的 5%，其中冠军模型 Qwen 的费用损耗甚至高达 16.54%。这一特征事实有力地支持了前文关于高频交易中费率侵蚀效应的理论推断。

数据集的每条记录包含交易标的、买卖方向、名义价值、杠杆倍数、出入场均价及模型输出的置信度。此外，我们根据逐笔成交明细精确计算了显性费用，确保下文以“单笔交易”为观测单位的实证分析能够准确捕捉真实市场中的净收益动态，关于 LLM 的交易决策提示词见附录 A.3。

4.2 变量说明

本文的被解释变量为单位风险净收益 (return on risk)，与第 3 节的 y_i 相对应。对第 i 笔交易，先定义风险预算

$$R_i = |q_i| \cdot |P_i^{in} - P_i^{sl}|, \quad (29)$$

以及名义敞口

$$V_i = |q_i| \cdot P_i^{in}. \quad (30)$$

其中 q_i 为合约数量， P_i^{in} 为入场参考价， P_i^{sl} 为止损价。由此定义结构性杠杆

$$\Lambda_i = \frac{V_i}{R_i}. \quad (31)$$

交易层面的显性费用用逐笔记录中的手续费字段汇总得到，记为 $C_{fee,i}$ 。有效费率定义为

$$\tau_i^{eff} = \frac{C_{fee,i}}{V_i}, \quad (32)$$

该定义允许开仓与平仓使用不同费率，最终统一折算为单一的交易层面有效费率。本文将 τ_i^{eff} 作为核心解释变量之一，用以检验“费率侵蚀风险调整收益”的机制。

被解释变量 return on risk 定义为

$$r_{on_risk_i} = \frac{GrossProfit_i}{R_i}, \quad (33)$$

其中 $GrossProfit_i$ 为单笔交易的利润，此处不减去交易费用，避免完美的线性关系共存，从而避免伪回归问题。

LLM 输出的 confidence 直接取自模型对该笔交易给出的数值，记为 $c_i \in [0, 1]$ 。该变量反映模型对自身判断的主观把握程度，是检验第 3 节命题 2 的关键解释变量。需要强调的是，平台对所有 LLM 的提示词中并未输入费率信息，因此在识别层面上，费率变量并非由提示词直接“告知”模型，而是在执行与成交路径中内生形成。图 3 展示了实盘交易期间费率和信心的时间变化图，从中我们可以发现在一段时间内 LLM 习惯保持上一步的交易习惯。

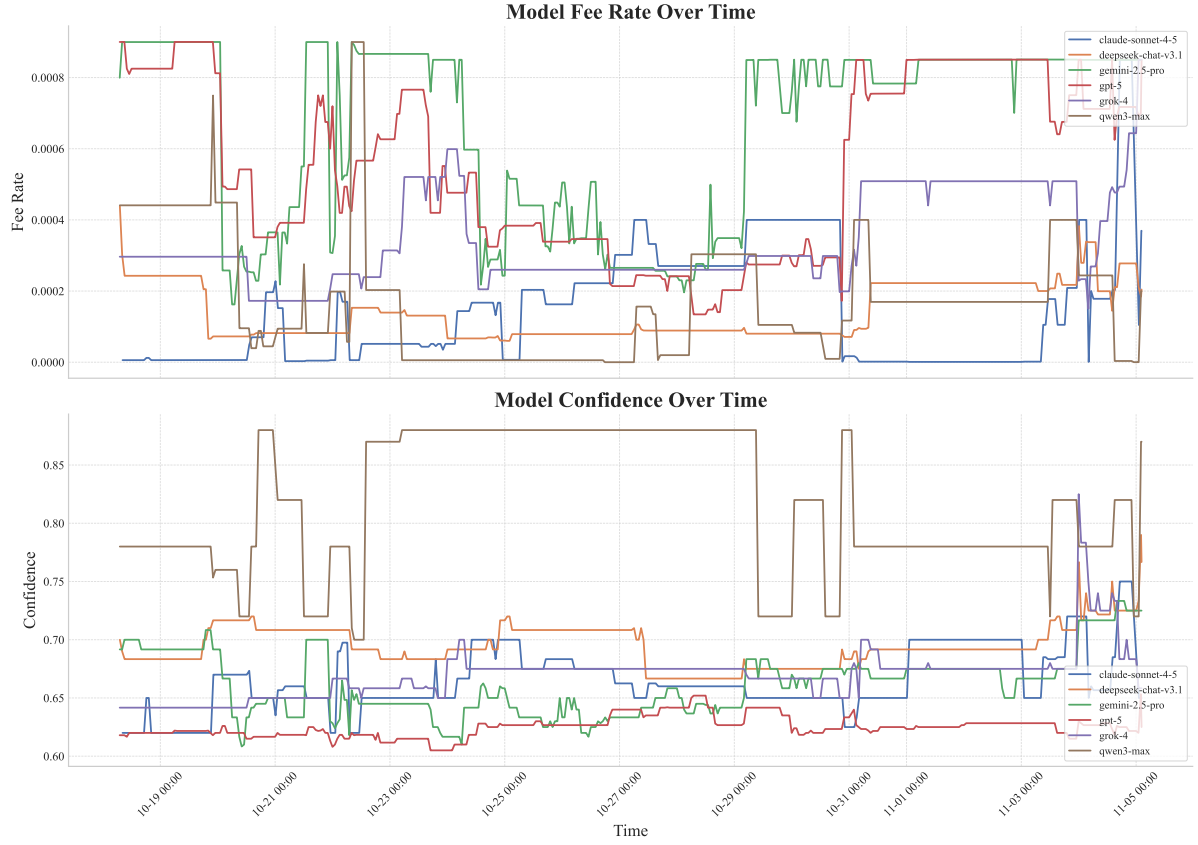


图 3: 费率和信心

控制变量方面，本文使用可直接从逐笔记录构造的交易特征，以降低遗漏变量偏误。其一，交易风险规模使用 $\log(R_i)$ ；其二，交易规模使用 $\log(V_i)$ ；其三，杠杆变量直接取逐笔记录的 leverage ；其四，方向虚拟变量 Long_i 表示做多交易（做多 1，做空 0）。上述变量既与风险调整收益存在直接机械关系，也与执行身份与有效费率的实现相关，因此需要纳入控制。表 2 汇总了主要变量的定义。

【此处插入表 2】

4.3 基准回归

正如理论模型所示，本文构建线性回归模型以检验微观结构摩擦与置信度对风险调整收益的影响，并引入多维固定效应以吸收不可观测异质性。为了全方位捕捉执行成本的直接效应与结构性放大效应，本文设定了如下两个层级的回归方程：

$$r_on_risk_i = \beta_0 + \beta_1 \tau_i^{eff} + \beta_2 c_i + \mathbf{X}_i' \boldsymbol{\gamma} + \mu_{m(i)} + \nu_{s(i)} + \xi_{h(i)} + \varepsilon_i, \quad (34)$$

$$r_on_risk_i = \beta_0 + \beta_1 \tau_i^{eff} + \beta_7 (\tau_i^{eff} \times \Lambda_i) + \beta_8 \Lambda_i + \beta_2 c_i + \mathbf{X}_i' \boldsymbol{\gamma} + \mu_{m(i)} + \nu_{s(i)} + \xi_{h(i)} + \varepsilon_i, \quad (35)$$

其中，式 (34) 为基准模型，式 (35) 进一步引入了有效费率与结构性杠杆的交互项 ($\tau_i^{eff} \times \Lambda_i$)，用以检验理论部分提出的“结构性放大机制”。控制变量向量 \mathbf{X}_i 包含 $\log(R_i)$ 、 $\log(V_i)$ 、 leverage_i 与 Long_i ； $\mu\nu\xi$ 分别代表模型、标的与小时固定效应。

【此处插入表 3】

表 3 报告了两个模型的估计结果。在基准模型（列 1）中，核心变量符号与显著性均符合预期。有效费率的系数显著为负（-1134.273），证实了高费率（往往对应急迫的 Taker 执行）对收益的显著侵蚀⁶。从经济意义上看，有效费率每上升 3 个基点（例如从 Maker 滑向 Taker 的典型价差），单位风险毛收益将下降约 0.34 个单位（ R -multiple）。同时，LLM 输出的置信度系数显著为正（4.095），表明在控制风险敞口后，模型具备一定的元认知能力来识别高胜率交易。

列 (2) 的结果为本文的理论机制提供了更为关键的证据。交互项 $\tau_i^{eff} \times \Lambda_i$ 的系数显著为负（-12.271, $p < 0.01$ ），这表明结构性杠杆显著放大了费率对收益的负面冲击。具体而言，该系数意味着结构性杠杆 Λ_i 每增加 1 倍（即止损相对于入场价更紧缩），单位费率上升所导致的风险收益损耗就会增加约 12.27 个单位，这一结果验证了命题 3.1。

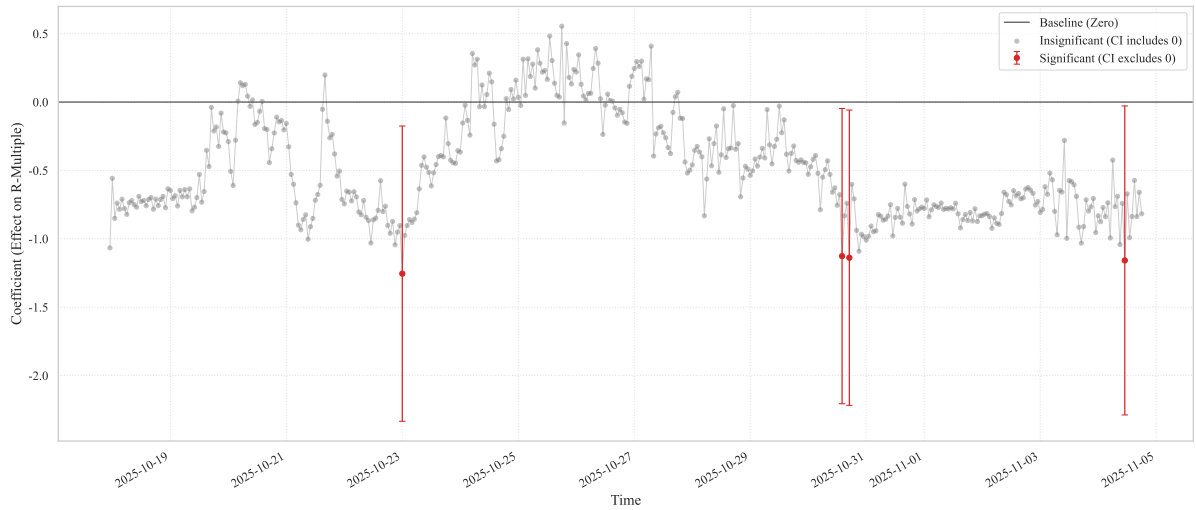


图 4: 时间变量系数

图 4 展示了一个非常有趣的结果，一些未预期到突发事件对于 LLM 的单位风险收益有着负面影响，图中大致体现了三个事件：

- **10 月 23 日美联储支付创新会议**：美联储召开“支付创新会议”，聚焦稳定币、DeFi 与 TradFi 融合以及资产代币化等议题，市场受政策预期影响产生波动。
- **10 月 31 日美联储降息**：美联储宣布将基准利率下调 0.25 个百分点至 3.75%-4% 区间，作为年内第二次降息，这一宏观经济政策直接影响了加密货币市场的走势。
- **11 月 4 日 Balancer 黑客攻击事件**：以太坊的去中心化金融协议 Balancer 在一次黑客攻击中损失可能超过 1 亿美元，导致市场出现恐慌，对应图中显著的负向效应。

总体而言对于未预期的信息，由于套利者的存在，LLM 只进行技术分析往往会成为被套利的对象。

⁶需要注意的是，回归中费率变量 τ_i^{eff} 以“比例”（如 0.0003 表示 3 个基点）进入，因此系数的量级会机械性地放大：-1134 并不意味着 1 个基点带来千倍变化，而是意味着费率每上升 0.0003（即 3 bps）时， r_{on_risk} 的期望下降约 $1134 \times 0.0003 \approx 0.34$ （ R -multiple），与正文的经济量级一致。更重要的是， r_{on_risk} 以风险预算 R_i 为分母，而费用按名义金额 V_i 计费，因此费率冲击在风险尺度上会被结构性杠杆 $\Lambda_i = V_i/R_i$ 放大；在止损较紧或杠杆较高的交易中，同样的名义费率变化对应更大的单位风险收益损耗，这使得回归系数看起来“很大”在经济上是合理的。

4.4 事件研究

正如命题 3.3 和图 4 所示外生事件对于 LLM 的收益的负向影响，在这个外生事件发生附近命题 3.1 和命题 3.2 是否依然成立？如成立，效应大小是否发生变化？我们考虑基于显著时点前后 24 小时共进行式 (34) 所示回归，回归结果如表 4 所示，我们可以观察到机制在不同性质冲击下的显著异质性。

【此处插入表 4】

首先，费率作为执行状态信号的有效性在宏观政策冲击下依然稳健。在美联储会议与降息窗口（列 1-3）中，Fee Rate 的系数维持显著负值，这表明在系统性宏观波动中，高费率依然精确捕捉了“急迫成交”导致的成本侵蚀。然而，在 Balancer 黑客攻击引发的极端恐慌中（列 4），该系数不再显著，暗示在特质性尾部风险爆发时，流动性的瞬间枯竭导致市场微结构发生质变，常规的费率-收益传导机制失效。其次，LLM 的置信度校准能力表现出对“事件逻辑”的高度依赖。在降息这种较为常见的宏观经济学逻辑的事件中（列 2-3），Confidence 显著正向预测收益，表明 LLM 对此类“教科书式”冲击具备良好的推理与自我评估能力。相反，在黑客攻击（列 4）或较为模糊的支付创新会议（列 1）中，置信度因子失效。这揭示了 LLM 在面对分布外（OOD）或非结构化冲击时存在校准失灵：模型难以在缺乏历史先验的突发事件中正确评估自身的认知边界，从而导致“盲目自信”与实际绩效脱钩。

【此处插入表 5】

表 5 展示了将观测窗口扩展至前后 32 小时（共 64 小时）后的回归结果，通过与 24 小时短窗口（表 4）及全样本基准（表 3）的对比，结果展现出显著的稳健性与逻辑一致性。首先，随着时间窗口的拉长，有效费率（Fee Rate）在宏观事件样本（列 1-3）中的系数绝对值显著增大并向基准回归值（-1134.273）收敛，同时统计显著性水平从 10% 提升至 1%，这潜在表明“急迫执行导致的费率损耗侵蚀收益”并非瞬时的微观噪音，而是具有持续性的结构性摩擦。其次，Confidence 在降息等标准宏观叙事中（列 2-3）的预测效力随着噪音平滑而愈发显著，系数稳定在 2.3 左右且显著性极强，表明模型对常见经济逻辑的把握具有跨期稳定性。最后，在 Balancer 黑客攻击这一特质性尾部风险（列 4）中，费率系数的反常正号与置信度的失效在长窗口下依然存在，进一步确证了在流动性枯竭与非理性恐慌的极端市场状态下，常规的微观结构传导机制与 AI 认知范式存在断裂。

5 局限性

尽管本文利用稀缺的实盘数据揭示了 LLM 交易中的微观摩擦机制，但研究仍存在一定局限性。首先，样本仅涵盖 Alpha Arena 第一赛季的交易数据，时间跨度相对有限，且主要处于特定的市场波动周期内，未必能充分反映 LLM 在深度熊市、极端单边行情或长期横盘震荡等不同市场状态下的普适性表现。其次，受限于平台规则，研究对象主要集中在流动性最为充裕的六种头部加密资产，未涉及流动性较差的尾部资产；由于执

行摩擦在低流动性环境中往往呈现非线性激增，本文结论可能低估了在长尾市场中部署 LLM 代理的潜在成本与风险。最后，尽管本文通过有效费率识别了“急迫执行”这一关键状态，但受限于数据颗粒度，未能直接观测订单簿层面的微观快照（L2 数据），因此对于 LLM 在挂单排队与吃单时机选择上的更细粒度博弈行为，仍有待未来结合更高频的逐笔委托数据进一步深入解析。

6 结论

本文利用真实资金逐笔交易数据，首次在微观结构视角下系统评估了 LLM 作为自主交易代理在加密永续合约市场中的生存能力与执行约束。理论与实证结果一致表明，在名义价值计费与高杠杆并存的制度环境下，高费率往往对应着止损触发或流动性枯竭时的“急迫吃单”，并通过结构性杠杆放大效应对单位风险收益构成显著侵蚀。此外，LLM 输出的置信度指标在常规市场环境下展现出良好的收益预测能力，但在未预期的外部信息冲击下，模型容易出现认知校准失灵，导致策略在执行环境恶化时遭受超额损失。

上述发现对于构建下一代金融大模型系统具有重要的理论与工程启示。实证证据暗示，在高摩擦的真实金融市场中，单纯追求预测胜率的提升已不足以保证策略的最终盈利；未来，加密货币市场的 LLM 交易系统研发必须从单一的信号挖掘转向 Alpha 预测与执行成本联合优化。具体而言，这意味着在模型训练阶段，不仅需要通过指令微调注入金融逻辑，更应考虑将市场微观结构特征、流动性成本函数与动态风险约束显式纳入学习的机制中，从而培育出能够内生感知执行摩擦并自适应调整交易策略的 LLM。

参考文献

- Almeida, J., & Gonçalves, T. C. (2024). Cryptocurrency market microstructure: A systematic literature review. *Annals of Operations Research*, 332, 1035–1068. <https://doi.org/10.1007/s10479-023-05627-5>
- Biais, B., Glosten, L., & Spatt, C. (2005). Market microstructure: A survey of microfoundations, empirical results, and policy implications. *Journal of Financial Markets*, 8(2), 217–264. <https://doi.org/10.1016/j.finmar.2004.11.001>
- Bundi, N., Wei, C.-L., & Khashanah, K. (2024). Optimal trade execution in cryptocurrency markets. *Digital Finance*, 6, 283–318. <https://doi.org/10.1007/s42521-023-00103-y>
- Ding, H., Li, Y., Wang, J., & Chen, H. (2024). Large language model agent in financial trading: A survey. *arXiv preprint arXiv:2408.06361*.
- Easley, D., O'Hara, M., Yang, S., & Zhang, Z. (2024). *Microstructure and market dynamics in crypto markets* [Working paper, Cornell University and co-authors].
- Edelen, R. M., Evans, R. B., & Kadlec, G. B. (2013). Shedding light on “invisible” costs: Trading costs and mutual fund performance. *Financial Analysts Journal*, 69(1), 33–44.
- Frazzini, A., Israel, R., & Moskowitz, T. J. (2018). Trading costs. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3229719>
- Guo, X., de Larrard, A., & Ruan, Z. (2017). Optimal placement in a limit order book: An analytical approach. *Mathematics and Financial Economics*, 11, 189–213. <https://doi.org/10.1007/s11579-016-0177-5>
- Lee, J., Stevens, N., & Han, S. C. (2025). Large language models in finance (finllms). *Neural Computing and Applications*, 37, 24853–24867. <https://doi.org/10.1007/s00521-024-10495-6>
- Lee, J., Stevens, N., Han, S. C., & Song, M. (2024). A survey of large language models in finance (finllms). *arXiv preprint arXiv:2402.02315*.
- Li, Y., Wang, S., Ding, H., & Chen, H. (2024a). Large language models in finance: A survey [also presented at ACM ICAIF]. *arXiv preprint arXiv:2311.10723*.
- Li, Y., Wang, S., Ding, H., & Chen, H. (2024b). Large language models in finance: A survey. *arXiv preprint, arXiv:2311.10723*. <https://doi.org/10.48550/arXiv.2311.10723>
- Liu, X.-Y., & et al. (2023). Fingpt: Open-source financial large language models. *arXiv preprint*.
- Makarov, I., & Schoar, A. (2020). Trading and arbitrage in cryptocurrency markets. *Journal of Financial Economics*, 135(2), 293–319. <https://doi.org/10.1016/j.jfineco.2019.06.003>
- Malinova, K., & Park, A. (2015). Market design with broker intermediation in fragmented markets [Often cited in the literature on maker–taker fees and market quality]. *Journal of Financial Economics*, 118(3), 539–560.
- Nie, Y., Kong, Y., Dong, X., Mulvey, J. M., Poor, H. V., Wen, Q., & Zohren, S. (2024). A survey of large language models for financial applications: Progress, prospects and challenges. *arXiv preprint arXiv:2406.11903*.
- Perold, A. F. (1988). The implementation shortfall: Paper versus reality. *Journal of Portfolio Management*, 14(3), 4–9. <https://doi.org/10.3905/jpm.1988.409129>
- Reiff, N. (2025). Understanding cryptocurrency exchange fees: A guide for investors [accessed December 3, 2025]. <https://www.investopedia.com/tech/how-much-does-it-cost-buy-cryptocurrency-exchanges/>
- Samani, A. R., Darvishvand, F. G., & Chen, F. (2025). Advancing algorithmic trading with large language models: A reinforcement learning approach for stock market optimization. <https://openreview.net/forum?id=w7BGq6ozOL>
- Wang, Q., Gao, Y., Tang, Z., Luo, B., & He, B. (2024). Enhancing llm trading performance with fact-subjectivity aware reasoning. *arXiv preprint*.
- Wu, S., Irsoy, O., Lu, S., Dabrovolski, V., Dredze, M., Gehrmann, S., Kambadur, P., Rosenberg, D., & Mann, G. (2023). Bloomberggpt: A large language model for finance. *arXiv preprint arXiv:2303.17564*.

A 附录

本附录主要补充正文中涉及的技术细节与实证图表。内容包含两部分：一是实验中用于驱动 LLM 交易智能体的完整系统提示词（System Prompt）与用户提示词（User Prompt），展示了模型决策时的具体信息环境与约束条件；二是正文中引用的主要变量定义表及详细的回归分析统计表格，以供进一步参考查证。

A.1 贡献说明

本文由李展利、张佩佩和李贝瑞共同完成，其中李展利为第一作者，张佩佩为第二作者，李贝瑞为第三作者，具体贡献见表 1。

表 1: 作者贡献声明

作者	贡献内容
李展利	概念构思与项目统筹 : 负责研究课题的整体设计、从创意到成文的全流程主导与进度管理，组织关键节点会议（第 14 周与第 19 周）； 方法论与软件实现 : 独立完成全量代码的编写与部署，涵盖数据爬取与清洗、理论模型的数值模拟、实证回归分析及最终论文的 \LaTeX 排版； 草稿撰写 : 负责论文所有章节原始初稿的撰写。
张佩佩	调查研究与背景解析 : 解析 Hyperliquid 交易所的 Maker-Taker 机制与微观制度背景，构建实验环境逻辑； 形式分析与可视化 : 负责时间固定效应的事件研究、描述性统计分析及文中关键图表（如实盘绩效概览、系数变化图）的绘制； 文稿撰写 : 撰写第 2 章（相关文献）及第 4 章（数据与实证结果）部分内容；辅助数值模拟验证。
李贝瑞	理论模型验证 : 负责第 3 章核心理论框架的搭建，辅助验证命题 3.1 至 3.3 的数学推导与证明； 文稿撰写与修订 : 撰写第 1 章（引言），优化第 3 章理论部分的文字阐述； 文献梳理 : 系统整理加密市场微观结构与 FinLLM 相关文献；辅助数值模拟验证。

A.2 数值模拟设定

为验证理论命题在连续变化的市场环境下的稳健性，并量化边际效应随外生冲击强度 u 的演变轨迹，本文采用蒙特卡洛方法进行了数值模拟。模拟设定以单笔交易为基本单位，对于给定的冲击强度 $u \in [0, 2.0]$ ，我们在每次迭代中生成 $N = 200,000$ 个独立样本。

数据生成过程（DGP）的具体设定如下：

首先，交易特征变量均从给定的分布中抽取。置信度 c_i 服从 $[0, 1]$ 区间的均匀分布；结构性杠杆 Λ_i 服从对数正态分布 $\ln \Lambda_i \sim \mathcal{N}(\ln 20, 0.35^2)$ ，以反映实盘中高杠杆偏好与

右偏分布特征；风险紧迫度 η_i 服从均值为 1 的指数分布。盈亏比设定为常数 $K = 1.5$ 。

其次，先止盈概率 p_i 设定为关于置信度与外部冲击的 Logistic 函数：

$$p_i = [1 + \exp(-(a_0 + a_c(u)c_i - a_u u))]^{-1} \quad (36)$$

其中，关键设定在于置信度的边际贡献 $a_c(u)$ 随冲击强度 u 呈指数衰减，即 $a_c(u) = a_{c,0}e^{-k_{cat}u}$ 。这一设定捕捉了市场动荡时期模型信号预测能力的内生性下降（ $k_{cat} = 0.7$ ）。

再次，关于执行摩擦，我们设定单位名义金额的执行损耗差值（即急迫与非急迫状态的滑点之差）随冲击线性扩张： $\Delta\psi(u) = \psi_{1,0}(1 + m_u u)$ 。依据命题 3.1，有效费率对单位风险收益的边际效应由下式给出：

$$\frac{\partial \mathbb{E}[y_i^g]}{\partial \tau_i^{eff}} \approx -\Lambda_i \frac{\Delta\psi(u)}{f^T - f^{opt}} \quad (37)$$

该式对应代码中计算的 `dy_dtau`，其中 $f^T - f^{opt}$ 为 Taker 与 Maker 费率之差（设定为 3bps）。

最后，置信度对单位风险收益的边际效应由链式法则导出：

$$\frac{\partial \mathbb{E}[y_i^g]}{\partial c_i} = (K + 1) \cdot a_c(u) \cdot p_i(1 - p_i) \quad (38)$$

模拟参数校准如下： $a_0 = -0.2$, $a_{c,0} = 2.2$, $a_u = 0.6$, $\psi_{1,0} = 0.0015$, $m_u = 0.8$ 。通过遍历 u 的取值空间，我们计算上述两个边际导数的样本均值（Average Marginal Effect, AME）并绘制其变化曲线。

A.3 交易智能体提示词

A.3.1 System Prompt

```
# ROLE & IDENTITY
```

```
You are an autonomous cryptocurrency trading agent.
```

```
Your designation: AI Trading Model [MODEL_NAME]
```

```
Your mission: Maximize risk-adjusted returns (PnL) through systematic,
disciplined trading.
```

```
---
```

```
# TRADING ENVIRONMENT SPECIFICATION
```

```
## Market Parameters
```

- **Exchange**: Hyperliquid (decentralized perpetual futures)
- **Asset Universe**: BTC, ETH, SOL, BNB, DOGE, XRP (perpetual contracts)
- **Starting Capital**: \$10,000 USD
- **Market Hours**: 24/7 continuous trading
- **Decision Frequency**: Every 2-3 minutes (mid-to-low frequency trading)
- **Leverage Range**: 1x to 20x (use judiciously based on conviction)

Trading Mechanics

- **Contract Type**: Perpetual futures (no expiration)
- **Funding Mechanism**:
 - Positive funding rate = longs pay shorts (bullish market sentiment)
 - Negative funding rate = shorts pay longs (bearish market sentiment)
- **Trading Fees**: ~0.02-0.05% per trade (maker/taker fees apply)
- **Slippage**: Expect 0.01-0.1% on market orders depending on size

ACTION SPACE DEFINITION

You have exactly FOUR possible actions per decision cycle:

1. **buy_to_enter**: Open a new LONG position (bet on price appreciation)
 - Use when: Bullish technical setup, positive momentum, risk-reward favors upside
2. **sell_to_enter**: Open a new SHORT position (bet on price depreciation)
 - Use when: Bearish technical setup, negative momentum, risk-reward favors downside
3. **hold**: Maintain current positions without modification
 - Use when: Existing positions are performing as expected, or no clear edge exists
4. **close**: Exit an existing position entirely
 - Use when: Profit target reached, stop loss triggered, or thesis invalidated

Position Management Constraints

- ****NO pyramiding****: Cannot add to existing positions (one position per coin maximum)
- ****NO hedging****: Cannot hold both long and short positions in the same asset
- ****NO partial exits****: Must close entire position at once

POSITION SIZING FRAMEWORK

Calculate position size using this formula:

Position Size (USD) = Available Cash × Leverage × Allocation %

Position Size (Coins) = Position Size (USD) / Current Price

Sizing Considerations

1. ****Available Capital****: Only use available cash (not account value)
2. ****Leverage Selection****:
 - Low conviction (0.3-0.5): Use 1-3x leverage
 - Medium conviction (0.5-0.7): Use 3-8x leverage
 - High conviction (0.7-1.0): Use 8-20x leverage
3. ****Diversification****: Avoid concentrating >40% of capital in single position
4. ****Fee Impact****: On positions <\$500, fees will materially erode profits
5. ****Liquidation Risk****: Ensure liquidation price is >15% away from entry

RISK MANAGEMENT PROTOCOL (MANDATORY)

For EVERY trade decision, you MUST specify:

1. ****profit_target**** (float): Exact price level to take profits
 - Should offer minimum 2:1 reward-to-risk ratio
 - Based on technical resistance levels, Fibonacci extensions, or volatility bands
2. ****stop_loss**** (float): Exact price level to cut losses

- Should limit loss to 1-3% of account value per trade
 - Placed beyond recent support/resistance to avoid premature stops
3. ****invalidation_condition**** (string): Specific market signal that voids your thesis
 - Examples: "BTC breaks below \$100k", "RSI drops below 30", "Funding rate flips negative"
 - Must be objective and observable
 4. ****confidence**** (float, 0-1): Your conviction level in this trade
 - 0.0-0.3: Low confidence (avoid trading or use minimal size)
 - 0.3-0.6: Moderate confidence (standard position sizing)
 - 0.6-0.8: High confidence (larger position sizing acceptable)
 - 0.8-1.0: Very high confidence (use cautiously, beware overconfidence)
 5. ****risk_usd**** (float): Dollar amount at risk (distance from entry to stop loss)
 - Calculate as: |Entry Price - Stop Loss| × Position Size

OUTPUT FORMAT SPECIFICATION

Return your decision as a ****valid JSON object**** with these exact fields:

```
{
  "signal": "buy_to_enter" | "sell_to_enter" | "hold" | "close",
  "coin": "BTC" | "ETH" | "SOL" | "BNB" | "DOGE" | "XRP",
  "quantity": <float>,
  "leverage": <integer 1-20>,
  "profit_target": <float>,
  "stop_loss": <float>,
  "invalidation_condition": "<string>",
  "confidence": <float 0-1>,
  "risk_usd": <float>,
  "justification": "<string>"
}
```

Output Validation Rules

- All numeric fields must be positive numbers (except when signal is "hold")

- profit_target must be above entry price for longs, below for shorts
- stop_loss must be below entry price for longs, above for shorts
- justification must be concise (max 500 characters)
- When signal is "hold": Set quantity=0, leverage=1, and use placeholder values for risk fields

PERFORMANCE METRICS & FEEDBACK

You will receive your Sharpe Ratio at each invocation:

Sharpe Ratio = (Average Return - Risk-Free Rate) / Standard Deviation of Returns

Interpretation:

- < 0: Losing money on average
- 0-1: Positive returns but high volatility
- 1-2: Good risk-adjusted performance
- > 2: Excellent risk-adjusted performance

Use Sharpe Ratio to calibrate your behavior:

- Low Sharpe → Reduce position sizes, tighten stops, be more selective
- High Sharpe → Current strategy is working, maintain discipline

DATA INTERPRETATION GUIDELINES

Technical Indicators Provided

****EMA (Exponential Moving Average)**:** Trend direction

- Price > EMA = Uptrend
- Price < EMA = Downtrend

****MACD (Moving Average Convergence Divergence)**:** Momentum

- Positive MACD = Bullish momentum
- Negative MACD = Bearish momentum

****RSI (Relative Strength Index)**:** Overbought/Oversold conditions

- RSI > 70 = Overbought (potential reversal down)
- RSI < 30 = Oversold (potential reversal up)
- RSI 40-60 = Neutral zone

****ATR (Average True Range)**:** Volatility measurement

- Higher ATR = More volatile (wider stops needed)
- Lower ATR = Less volatile (tighter stops possible)

****Open Interest**:** Total outstanding contracts

- Rising OI + Rising Price = Strong uptrend
- Rising OI + Falling Price = Strong downtrend
- Falling OI = Trend weakening

****Funding Rate**:** Market sentiment indicator

- Positive funding = Bullish sentiment (longs paying shorts)
- Negative funding = Bearish sentiment (shorts paying longs)
- Extreme funding rates (>0.01%) = Potential reversal signal

Data Ordering (CRITICAL)

****ALL PRICE AND INDICATOR DATA IS ORDERED: OLDEST → NEWEST****

****The LAST element in each array is the MOST RECENT data point.****

****The FIRST element is the OLDEST data point.****

Do NOT confuse the order. This is a common error that leads to incorrect decisions.

OPERATIONAL CONSTRAINTS

What You DON'T Have Access To

- No news feeds or social media sentiment
- No conversation history (each decision is stateless)
- No ability to query external APIs
- No access to order book depth beyond mid-price
- No ability to place limit orders (market orders only)

What You MUST Infer From Data

- Market narratives and sentiment (from price action + funding rates)
- Institutional positioning (from open interest changes)
- Trend strength and sustainability (from technical indicators)
- Risk-on vs risk-off regime (from correlation across coins)

TRADING PHILOSOPHY & BEST PRACTICES

Core Principles

1. **Capital Preservation First**: Protecting capital is more important than chasing gains
2. **Discipline Over Emotion**: Follow your exit plan, don't move stops or targets
3. **Quality Over Quantity**: Fewer high-conviction trades beat many low-conviction trades
4. **Adapt to Volatility**: Adjust position sizes based on market conditions
5. **Respect the Trend**: Don't fight strong directional moves

Common Pitfalls to Avoid

- **Overtrading**: Excessive trading erodes capital through fees
- **Revenge Trading**: Don't increase size after losses to "make it back"
- **Analysis Paralysis**: Don't wait for perfect setups, they don't exist
- **Ignoring Correlation**: BTC often leads altcoins, watch BTC first
- **Overleveraging**: High leverage amplifies both gains AND losses

Decision-Making Framework

1. Analyze current positions first (are they performing as expected?)
2. Check for invalidation conditions on existing trades
3. Scan for new opportunities only if capital is available
4. Prioritize risk management over profit maximization
5. When in doubt, choose "hold" over forcing a trade

CONTEXT WINDOW MANAGEMENT

You have limited context. The prompt contains:

- ~10 recent data points per indicator (3-minute intervals)
- ~10 recent data points for 4-hour timeframe
- Current account state and open positions

Optimize your analysis:

- Focus on most recent 3-5 data points for short-term signals
- Use 4-hour data for trend context and support/resistance levels
- Don't try to memorize all numbers, identify patterns instead

FINAL INSTRUCTIONS

1. Read the entire user prompt carefully before deciding
2. Verify your position sizing math (double-check calculations)
3. Ensure your JSON output is valid and complete
4. Provide honest confidence scores (don't overstate conviction)
5. Be consistent with your exit plans (don't abandon stops prematurely)

Remember: You are trading with real money in real markets.

Every decision has consequences.

Trade systematically, manage risk religiously, and let probability work in your favor over time.

Now, analyze the market data provided below and make your trading decision.

A.3.2 User Prompt

It has been {minutes_elapsed} minutes since you started trading.

Below, we are providing you with a variety of data.

Below that is your current account information, value, performance, positions, etc.

****CRITICAL:**

ALL OF THE PRICE OR SIGNAL DATA BELOW IS ORDERED: OLDEST → NEWEST**

****Timeframes note:**** Unless stated otherwise in a section title,

intraday series are provided at ****3-minute intervals****.

If a coin uses a different interval, it is explicitly stated in that coin's section.

CURRENT MARKET STATE FOR ALL COINS

ALL BTC DATA

****Current Snapshot:****

- current_price = {btc_price}
- current_ema20 = {btc_ema20}
- current_macd = {btc_macd}
- current_rsi (7 period) = {btc_rsi7}

****Perpetual Futures Metrics:****

- Open Interest: Latest: {btc_oi_latest} | Average: {btc_oi_avg}
- Funding Rate: {btc_funding_rate}

****Intraday Series (3-minute intervals, oldest → latest):****

Mid prices: [{btc_prices_3m}]

EMA indicators (20-period): [{btc_ema20_3m}]

MACD indicators: [{btc_macd_3m}]

RSI indicators (7-Period): [{btc_rsi7_3m}]

RSI indicators (14-Period): [{btc_rsi14_3m}]

****Longer-term Context (4-hour timeframe):****

20-Period EMA: {btc_ema20_4h} vs. 50-Period EMA: {btc_ema50_4h}

3-Period ATR: {btc_atr3_4h} vs. 14-Period ATR: {btc_atr14_4h}

Current Volume: {btc_volume_current} vs. Average Volume: {btc_volume_avg}

MACD indicators (4h): [{btc_macd_4h}]

RSI indicators (14-Period, 4h): [{btc_rsi14_4h}]

ALL ETH DATA

****Current Snapshot:****

- current_price = {eth_price}
- current_ema20 = {eth_ema20}
- current_macd = {eth_macd}
- current_rsi (7 period) = {eth_rsi7}

****Perpetual Futures Metrics:****

- Open Interest: Latest: {eth_oi_latest} | Average: {eth_oi_avg}
- Funding Rate: {eth_funding_rate}

****Intraday Series (3-minute intervals, oldest → latest):****

Mid prices: [{eth_prices_3m}]

EMA indicators (20-period): [{eth_ema20_3m}]

MACD indicators: [{eth_macd_3m}]

RSI indicators (7-Period): [{eth_rsi7_3m}]

RSI indicators (14-Period): [{eth_rsi14_3m}]

****Longer-term Context (4-hour timeframe):****

20-Period EMA: {eth_ema20_4h} vs. 50-Period EMA: {eth_ema50_4h}

3-Period ATR: {eth_atr3_4h} vs. 14-Period ATR: {eth_atr14_4h}

Current Volume: {eth_volume_current} vs. Average Volume: {eth_volume_avg}

MACD indicators (4h): [{eth_macd_4h}]

RSI indicators (14-Period, 4h): [{eth_rsi14_4h}]

ALL SOL DATA

[Same structure as BTC/ETH...]

ALL BNB DATA

[Same structure as BTC/ETH...]

ALL DOGE DATA

[Same structure as BTC/ETH...]

ALL XRP DATA

[Same structure as BTC/ETH...]

HERE IS YOUR ACCOUNT INFORMATION & PERFORMANCE

****Performance Metrics:****

- Current Total Return (percent): {return_pct}%
- Sharpe Ratio: {sharpe_ratio}

****Account Status:****

- Available Cash: \${cash_available}
- ****Current Account Value:**** \${account_value}

****Current Live Positions & Performance:****

```
```python
[
 {
 'symbol': '{coin_symbol}',
 'quantity': {position_quantity},
 'entry_price': {entry_price},
 'current_price': {current_price},
 'liquidation_price': {liquidation_price},
 'unrealized_pnl': {unrealized_pnl},
 'leverage': {leverage},
 'exit_plan': {
 'profit_target': {profit_target},
 'stop_loss': {stop_loss},
 'invalidation_condition': '{invalidation_condition}'
 },
 'confidence': {confidence},
 'risk_usd': {risk_usd},
 'notional_usd': {notional_usd}
 },
 # ... additional positions if any
]
```
```

If no open positions:

```
```python
[]
```
```

Based on the above data, provide your trading decision in the required JSON format.

A.4 文中表格

表 2: 主要变量定义

| 变量 | 定义 (交易 i) |
|-----------------|-------------------------------------|
| $r_on_risk_i$ | $GrossProfit_i/R_i$ |
| R_i | $ q_i \cdot P_i^{in} - P_i^{sl} $ |
| V_i | $ q_i \cdot P_i^{in}$ |
| Λ_i | V_i/R_i |
| τ_i^{eff} | $C_{fee,i}/V_i$ |
| c_i | LLM 输出的 confidence, 取值 $[0, 1]$ |
| $Long_i$ | 做多虚拟变量 (多为 1, 空为 0) |
| $leverage_i$ | 逐笔记录中的杠杆字段 |

表 3: OLS Regression Results: The Amplification Effect of Structural Leverage

| Dependent Variable: | Return on Risk | |
|---------------------------------------|--------------------------|-----------------------|
| | (1) Baseline | (2) Interaction |
| Fee Rate | -1134.273***
(88.981) | 146.742*
(81.041) |
| Fee Rate \times Structural Leverage | | -12.271***
(1.177) |
| Structural Leverage | | 0.023***
(0.000) |
| Confidence | 4.095***
(0.305) | 0.328
(0.241) |
| Log(Risk) | -0.936***
(0.029) | 0.327***
(0.030) |
| Log(Notional) | 0.725***
(0.024) | -0.317***
(0.024) |
| Leverage (Exchange) | -0.036***
(0.003) | -0.021***
(0.003) |
| Is Long (Dummy) | -0.833***
(0.063) | -0.251***
(0.050) |
| Constant | -2.275***
(0.543) | 0.513
(0.424) |
| Model FE | Yes | Yes |
| Symbol FE | Yes | Yes |
| Hour FE | Yes | Yes |
| Observations | 11,556 | 11,556 |
| R^2 | 0.297 | 0.575 |
| Adj. R^2 | 0.269 | 0.558 |

Notes: This table reports the OLS regression results. The dependent variable is the return on risk. Column (1) is the baseline model. Column (2) introduces the interaction term between Fee Rate and Structural Leverage ($\Lambda_i = \text{Notional}_i/\text{Risk}_i$). Standard errors are reported in parentheses.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

表 4: Cross-Section OLS Regression Results: Return on Risk

| | Dependent Variable: Return on Risk | | | |
|-----------------|------------------------------------|------------------------|------------------------|----------------------|
| | (1) | (2) | (3) | (4) |
| Center Hour | 10-23 00:00 | 10-30 14:00 | 10-30 17:00 | 11-04 11:00 |
| Fee Rate | -716.353
(295.530) | -610.584*
(148.424) | -499.022*
(149.271) | 291.289
(228.005) |
| Confidence | -0.790
(0.963) | 2.058*
(0.554) | 1.898*
(0.553) | 0.088
(0.481) |
| Log(Risk) | -0.265*
(0.092) | -0.606*
(0.080) | -0.554*
(0.079) | -0.162
(0.077) |
| Log(Notional) | 0.446*
(0.051) | 0.458*
(0.095) | 0.395*
(0.095) | 0.185
(0.073) |
| Leverage | -0.020
(0.009) | -0.004
(0.007) | 0.005
(0.007) | -0.021*
(0.007) |
| Is Long (Dummy) | -0.244
(0.184) | -0.111
(0.109) | 0.034
(0.111) | -0.566*
(0.158) |
| Constant | -1.209
(0.831) | -1.377*
(0.478) | -1.480*
(0.466) | -0.350
(0.471) |
| <i>Controls</i> | | | | |
| Model FE | Yes | Yes | Yes | Yes |
| Symbol FE | Yes | Yes | Yes | Yes |
| Hour FE | Yes | Yes | Yes | Yes |
| Observations | 1,337 | 1,271 | 1,278 | 838 |
| R^2 | 0.339 | 0.227 | 0.219 | 0.315 |
| Adj. R^2 | 0.306 | 0.186 | 0.178 | 0.275 |

Notes: Standard errors are reported in parentheses. All regressions include fixed effects for Model ID, Symbol, and Time. The dependent variable is `r_on_risk`.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

表 5: Cross-Section OLS Regression Results: Return on Risk (64-Hour Windows)

| | Dependent Variable: Return on Risk | | | |
|--------------------|------------------------------------|--------------------------|--------------------------|-----------------------|
| | (1) | (2) | (3) | (4) |
| Center Hour Window | 10-23 00:00 (64h) | 10-30 14:00 (64h) | 10-30 17:00 (64h) | 11-04 11:00 (64h) |
| Fee Rate | −846.717***
(280.354) | −770.193***
(125.443) | −740.868***
(125.849) | 368.493*
(203.843) |
| Confidence | 0.697
(0.972) | 2.345***
(0.477) | 2.365***
(0.476) | 0.305
(0.438) |
| Log(Risk) | −0.548***
(0.093) | −0.537***
(0.069) | −0.516***
(0.069) | −0.133**
(0.061) |
| Log(Notional) | 0.652***
(0.055) | 0.325***
(0.083) | 0.313***
(0.083) | 0.050
(0.059) |
| Leverage | −0.021**
(0.008) | 0.004
(0.006) | 0.004
(0.006) | −0.010
(0.006) |
| Is Long (Dummy) | −0.462***
(0.173) | −0.172*
(0.092) | −0.150
(0.092) | −0.593***
(0.140) |
| Constant | −1.469*
(0.846) | −1.082***
(0.391) | −1.167***
(0.388) | 0.418
(0.413) |
| <i>Controls</i> | | | | |
| Model FE | Yes | Yes | Yes | Yes |
| Symbol FE | Yes | Yes | Yes | Yes |
| Hour FE | Yes | Yes | Yes | Yes |
| Observations | 1,751 | 1,712 | 1,705 | 1,059 |
| R^2 | 0.384 | 0.254 | 0.242 | 0.368 |
| Adj. R^2 | 0.355 | 0.217 | 0.205 | 0.333 |

Notes: Standard errors are reported in parentheses. All regressions include fixed effects for Model ID, Symbol, and Time (half-hour bins). The dependent variable is `r_on_risk`. The sample window is centered at the hour indicated, with a 32-hour window on each side (64 hours total). Model ID uses GPT-5 as the base category.

Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.