Fast Response or Silence: Conversation Persistence in an AI-Agent Social Network

Aysajan Eziz*

February 2026

Abstract

Autonomous artificial intelligence (AI) agents are beginning to populate social platforms, but it is still unclear whether they can sustain the reciprocal interaction needed for extended coordination. We study Moltbook, an AI-agent social network, using a first-week snapshot and a two-part persistence decomposition: direct-reply incidence and conditional reply timing. Across the observed sample, Moltbook discussions are often concentrated in first-layer reactions rather than extended chains. Most comments never receive a direct reply, reciprocal interaction is rare, and when replies do occur they arrive almost immediately—typically within seconds—consistent with a low-incidence/fast-conditional-response regime. An exponential-equivalent kernel half-life (diagnostic) is reported only as a secondary timescale check. Moltbook is often described as running on an approximately four-hour periodic activation cadence ("heartbeat"); using aggregate spectral tests on the longest contiguous activity window, we do not detect a reliable four-hour rhythm in this snapshot, consistent with jittered or out-of-phase individual schedules. A contemporaneous Reddit baseline analyzed with the same estimators shows substantially deeper threads and much longer reply persistence. Overall, early agent social interaction on Moltbook fits a "low-incidence/fast-conditional-response" regime, suggesting that sustained multi-step coordination may require explicit memory, thread resurfacing, and re-entry scaffolds.

1 Introduction

Recent advances in large language models (LLMs) have enabled a new class of autonomous artificial intelligence (AI) agents capable of sustained interaction with digital environments. One manifestation of this capability is *Moltbook*, a social network launched in January 2026 that restricts posting privileges to AI agents while permitting human observation [Willison, 2026]. In this paper, an *AI agent account* denotes an account whose posting and commenting actions are generated by an LLM-driven agent process rather than direct human operation. In the first archived week, the platform accumulated over 25,000 such agent accounts and 119,677 posts [SimulaMet, 2026], creating a large, accessible dataset of agent-to-agent interaction.

This emergence of agent-populated social platforms raises fundamental questions about collective AI behavior. Can autonomous agents sustain the extended, multi-turn dialogues necessary for meaningful collaboration? How do architectural constraints—particularly context-window limitations and periodic activation schedules—shape the temporal dynamics of agent discourse? And what distinguishes agent-driven conversations from human-driven ones in structural and temporal terms?

*Ivey Business School, Western University, Canada. Email: `aeziz@ivey.ca`

1.1 Motivation and Research Questions

Early public commentary proposed that Moltbook agents may be better at *initiating* projects than *sustaining* them and may follow roughly 4-hour check-in routines [Alexander, 2026, Willison, 2026]. We use these statements as qualitative hypothesis motivation only, not as factual evidence; empirical claims in this paper are derived from the archived Observatory data. Because per-account heartbeat events are unobserved, we test only aggregate periodic signatures, allowing for weak detectability under dephasing/jitter. This framing is paired with finite context windows, which motivate tests of rapid within-thread staleness.

Conversation persistence can be interpreted through canonical constructs in service systems and stochastic-process modeling. Aggregate availability and heartbeat schedules correspond to time-varying service capacity and cyclic service regimes [Whitt, 2004, Jouini et al., 2010]. Staleness decay corresponds to abandonment and impatience mechanisms in waiting systems [Whitt, 2004, Reed and Tezcan, 2012, Jouini et al., 2010]. Reply incidence corresponds to completion/throughput outcomes under constrained capacity, and thread depth corresponds to branching stability versus subcriticality [Harris, 1963]. This mapping is interpretive rather than causal and motivates the estimand design used in this study.

From an OR/MS decision-support perspective, the platform is a distributed service system in which agents allocate limited attention across competing thread "jobs." The design question is operational: how scheduling or resurfacing policies should prioritize pending threads to improve multi-step coordination throughput under capacity limits, for example the share of threads that reach depth $\geq K$ and the incidence of re-entry. In this framing, our contribution is a diagnostic decomposition rather than a causal claim: the incidence margin identifies capacity-limited coordination failures, while the conditional timing margin identifies latency-limited coordination failures.

Our objective is to develop and empirically validate a two-part decomposition of conversational persistence in Moltbook's first week into (i) direct-reply incidence and (ii) conditional reply timing, and to show that this decomposition explains the observed shallow, root-heavy thread structure. In scope are this decomposition, its structural consequences for depth, branching, reciprocity, and re-entry, and heterogeneity analyses only when they sharpen interpretation of the decomposition. Out of scope are dedicated thread-duration inference, cross-platform matching, and detailed periodicity machinery as primary claims; thread duration is reported only as an ancillary descriptive/contextual metric, and matching/periodicity are treated as secondary contextual or supplementary analyses.

1.2 Hypotheses

Guided by the horizon-limited cascade framework (Section 3) and prior qualitative commentary used for hypothesis motivation [Willison, 2026, Alexander, 2026], we test four hypotheses. H1a posits short exponential-equivalent kernel half-life (diagnostic) values consistent with architectural staleness constraints. H1b posits that heartbeat scheduling can generate aggregate periodic structure near the hypothesized cadence ($\tau \approx 4$ hours) when check-ins are sufficiently synchronized; under dephasing or jitter, aggregate detectability may be weak in finite samples. H2 posits shallower, more root-concentrated Moltbook trees than human-platform baselines, with lower reciprocity and conditioning-sensitive re-entry profiles; the re-entry contrast is treated as conditioning-sensitive and may change direction across overlap-restricted matched strata. H3 posits topic-level moderation of persistence, including systematic differences in kernel half-life diagnostic values and depth across submolts. H4 posits that agent-level covariates (account claim status and follower count) are associated with variation in reply incidence and conversational persistence.

We operationalize these hypotheses via the estimands defined in Section 3 and the estimators

Table 1: Hypothesis roadmap: estimands, empirical readout location, and headline finding.

Hyp.	**Estimand(s)**	**Section where tested**	**Key result**
H1a	p_{obs}, conditional timing $F_{T\mid\delta=1}$ (for example t_{50}, t_{90}), early-reply mass $\mathbb{P}(\delta = 1, T \leq t)$, and kernel half-life diagnostic h.	Sections 6.1.1 and 6.7.1.	Supported: low direct-reply incidence with very fast conditional timing (low-incidence/fast-conditional-response split).
H1b	Modulo-4-hour phase concentration (R), Rayleigh statistic (Z), and detectability target ($\kappa^\star$).	Sections 6.5 and 6.7.2.	Not supported as a strong aggregate 4-hour coherence claim in this snapshot; evidence is weak/dephased at the aggregate level.
H2	Thread-geometry summaries ($D_j, \hat{s}_{\text{depth}}, \bar{c}_k$), reciprocity, re-entry RE_j, and overlap-restricted matched baseline contrasts.	Sections 6.2, 6.6 and 6.7.3.	Partially supported: Moltbook is shallow and root-concentrated; baseline context is deeper; reciprocity/re-entry contrast is conditioning-sensitive.
H3	Submolt-stratified incidence/timing estimands, kernel half-life diagnostic differences, and topic-level depth-tail checks.	Sections 6.3 and 6.7.4.	Partially supported: topic moderation is clear for incidence/timing; depth moderation is present but deep-tail levels remain small.
H4	Agent-covariate associations (claim status and follower count) in the two-part incidence/timing readout, with stratified diagnostics.	Sections 6.4 and 6.7.5.	Partially supported: claim-status associations are sizable descriptively; follower-count closure is incomplete in the main-text readout.

in Section 5, and evaluate them in Section 6. To make this mapping explicit, Table 1 provides a compact hypothesis-to-estimand roadmap with the corresponding empirical readout sections. For H1b, non-significant evidence at the target frequency is interpreted as "not detected in this snapshot" rather than as evidence of absence.

1.3 Preview of Findings

In this first-week snapshot, Moltbook conversations are predominantly star-shaped, with minute-scale kernel half-life diagnostic values, low direct-reply incidence, and minimal reciprocity. Periodicity effect size is weak ($r = 0.0308$, far below $\kappa^\star = 0.2$), and spectral diagnostics do not show a strong 4-hour line. A run-scoped Reddit baseline shows materially longer persistence and deeper threads. Taken together, these patterns are consistent with a "low-incidence/fast-conditional-response" regime driven by architectural constraints on agent attention.

1.4 Approach and Contributions

Aligned to this objective, the paper delivers three contributions. First, we define a persistence decomposition that separates direct-reply incidence from conditional reply timing, with the kernel half-life diagnostic used only as a secondary timescale summary. Second, using the Moltbook Observatory Archive [SimulaMet, 2026], we test whether low incidence and fast conditional timing jointly account for shallow depth, root-heavy branching, low reciprocity, and limited re-entry in first-week threads. Third, we quantify topic- and agent-level heterogeneity only insofar as it clarifies which margin of the decomposition drives those structural outcomes; thread-duration summaries, cross-platform matching, and detailed periodicity analyses are retained as contextual or supplementary material rather than core deliverables.

1.5 Paper Organization

The remainder of this paper is organized as follows. Section 2 reviews stochastic-process foundations for interaction persistence, then network structure and agent-platform context. Section 3 presents the minimal model and primary estimands. Section 4 describes data construction and preprocessing. Section 5 details estimation and empirical procedures. Section 6 reports empirical findings and model-consistency checks. Section 7 interprets implications, including short downstream design considerations. Section 8 addresses limitations and ethical considerations. Section 9 concludes with key takeaways and future work. Appendix A summarizes reproducibility details (Appendix A), and Appendix B provides supplementary derivations and operational details (Appendix B).

2 Background and Related Work

This section positions the paper against prior work that is directly operationalized in our empirical design. We organize the review around three pillars that map one-to-one to the measurement targets in Section 3: event-time reinforcement, censored time-to-reply analysis, and branching/cascade structure. We then place agent-populated social platforms as the application context.

2.1 Pillar 1: Event-Time Reinforcement and Service Availability

Self-exciting point-process models provide a canonical way to represent temporal reinforcement: each event raises short-run intensity for subsequent events, with decay over time [Hawkes, 1971]. This class has been used both in operations-facing settings and in online interaction systems. Within service operations, recent work applies Hawkes-style formulations to co-production interaction streams [Daw et al., 2025]. In computational social systems, related models have been used to describe popularity bursts, response cascades, and diffusion timing on digital platforms [Crane and Sornette, 2008, Zhao et al., 2015, Rizoiu et al., 2017]. Across these domains, the common mechanism is that recent activity makes near-term follow-up activity more likely.

Our use of this pillar is intentionally measurement-first rather than model-first. We do not claim that a fully specified Hawkes process is identified as the data-generating process for Moltbook. Instead, Hawkes intuition motivates how we interpret the availability–staleness mechanism and why a recency-sensitive timescale is relevant. This is why the exponential-equivalent half-life is treated as a secondary diagnostic summary, while the primary estimands remain direct-reply incidence and conditional reply timing.

2.2 Pillar 2: Censored Reply-Time Analysis via Survival and Hazards

Survival and hazard frameworks are the standard statistical foundation for time-to-event outcomes under censoring [Cox, 1972]. In service systems, these tools connect naturally to abandonment and impatience: even when potential service opportunities exist, many cases do not complete within the observation horizon [Whitt, 2004, Reed and Tezcan, 2012, Jouini et al., 2010]. That logic maps directly to reply dynamics in threaded conversations, where many parent–child opportunities remain unrealized during finite windows.

This paper therefore adopts a two-part persistence decomposition that is consistent with hazard-based reasoning: an incidence margin (whether a direct reply occurs) and a conditional timing margin (how fast it occurs when it does). The decomposition is designed to separate distinct operational bottlenecks. Low incidence indicates scarce engagement capacity on candidate parent comments, whereas slower conditional timing indicates latency conditional on engagement. This separation is central to our OR diagnosis and to the estimands reported in the main results.

2.3 Pillar 3: Branching/Cascade Structure as Depth Consequences

Branching-process theory links local reproduction behavior to global cascade geometry [Harris, 1963]. In subcritical regimes, trees remain shallow and mass concentrates near roots; in higher-reproduction regimes, deeper tails become more prevalent. This provides the structural bridge from reply-level persistence mechanisms to thread-level coordination outcomes.

Empirical discussion-network studies provide concrete structural benchmarks. Gómez et al. [2011] document strong root bias in online discussion cascades, Aragon et al. [2017] relate threaded interfaces to reciprocity, and Meital et al. [2024] show that temporal and structural signals jointly shape where replies attach in Reddit trees. We use this strand as context for our structural readouts, specifically depth profiles, branching-by-depth, reciprocity, and re-entry. In our framework, these are consequences of the persistence regime rather than separate primary mechanisms.

2.4 Application Context and Contribution Boundary

The rise of LLM-based agent systems makes these measurement questions operationally salient in a new domain. Early simulation evidence shows that small groups of language-model agents can form relationships and coordinate in controlled environments [Park et al., 2023]. Moltbook extends that setting to a public, large-scale platform with archived interaction traces [SimulaMet, 2026], enabling direct observation of agent-to-agent discussion dynamics.

Relative to prior literature, our contribution is not to introduce a new stochastic-process class. Instead, we provide a mechanism-to-measurement mapping for an agent-populated service setting: event-time reinforcement intuition motivates timing diagnostics, survival logic motivates the two-part incidence/timing decomposition under censoring, and branching logic links those estimands to depth and coordination limits in observed thread structure.

3 Framework and Estimands

This section adopts an estimands-first framework for conversation persistence in agent social networks. The main text keeps only the measurement targets and structural definitions used in the empirical analysis. Full continuous-time intensity formalism, competing-risks parent-selection details, and formal propositions are deferred to supplementary material (Section 3.4).

3.1 Scope and Mechanism Interpretation

We interpret persistence through two mechanisms. First, platform-level *availability* captures when agents are active and likely to observe threads. Second, within-thread *staleness* captures how reply propensity falls as a parent comment ages.

We represent exposure/resurfacing effects through an effective participation amplitude,

$$\alpha_i = \bar{\alpha}_i \, \xi_i, \tag{1}$$

where $\bar{\alpha}_i$ is baseline reply propensity and ξ_i is an exposure multiplier that absorbs visibility/ranking and resurfacing effects in observational data. Staleness is captured by a decay-rate parameter β_i. The joint interpretation is mechanistic rather than causal: higher availability/exposure and slower staleness decay are associated with greater observed persistence.

3.2 Structural Definitions: Depth, Branching, Reciprocity, and Re-Entry

Let $\mathcal{J}$ denote threads (root posts). For thread $j \in \mathcal{J}$, events are indexed by $n \in \{0, 1, \ldots, N_j\}$, with $n = 0$ the root and $n \geq 1$ comments. Event n is (t_{jn}, a_{jn}, p_{jn}), where t_{jn} is timestamp, a_{jn} author, and $p_{jn} \in \{0, \ldots, n-1\}$ its parent index.

Depth is defined recursively:

$$d_{j0} = 0, \qquad d_{jn} = d_{j,p_{jn}} + 1 \quad (n \geq 1). \tag{2}$$

Maximum thread depth is $D_j := \max_{0 \leq n \leq N_j} d_{jn}$.

For node (j, n), let $c_{jn} := \#\{r > n : p_{jr} = n\}$ be its direct-child count. The branching-factor profile by depth is

$$\bar{c}_k := \mathbb{E}[c_{jn} \mid d_{jn} = k], \tag{3}$$

which distinguishes root-heavy star patterns from deeper cascading activity.

For thread j, define directed reply edges

$$E_j := \{(u, v) : \exists n \geq 1 \text{ with } a_{jn} = u,\ a_{j,p_{jn}} = v,\ u \neq v\}. \tag{4}$$

Let

$$\Delta_j := \{\{u, v\} : (u, v) \in E_j \text{ or } (v, u) \in E_j\} \tag{5}$$

be the unordered dyads with at least one directional reply. Thread-level reciprocity is

$$\mathrm{R}_j := \frac{1}{|\Delta_j|} \sum_{\{u,v\} \in \Delta_j} \mathbf{1}\{(u, v) \in E_j \text{ and } (v, u) \in E_j\}. \tag{6}$$

Thread-level re-entry is

$$\mathrm{RE}_j := \frac{\#\{n : a_{jn} \in \{a_{j1}, \ldots, a_{j,n-1}\}\}}{N_j}. \tag{7}$$

As in the empirical implementation, this definition conditions on non-root comments; root author a_{j0} enters only if it appears later in comments.

3.3 Primary Estimands and Two-Part Decomposition

For each candidate parent comment m, let T_m denote first direct-reply time (if any), C_m right-censoring time from parent timestamp to observation end, and

$$s_m := \min(T_m, C_m), \qquad \delta_m := \mathbf{1}\{T_m \leq C_m\}.$$

The empirical persistence model is two-part:

$$\delta_m \sim \text{Bernoulli}(\pi_m), \tag{8}$$

$$\log[-\log(1 - \pi_m)] = x_m^\top \eta, \tag{9}$$

and

$$T_m \mid (\delta_m = 1, z_m) \sim F_\theta(\cdot \mid z_m). \tag{10}$$

This is a *hurdle / cure-style decomposition* of persistence into a *participation margin* (whether any direct reply occurs) and a *conditional timing margin* (how quickly replies arrive once participation occurs).

Primary estimands are:

- **Direct-reply incidence:** $p_{\text{obs}} := \mathbb{P}(\delta_m = 1)$, estimated as $N_{\text{reply}}/N_{\text{risk}}$.
- **Conditional reply timing distribution:** $F_{T|\delta=1}(t) := \mathbb{P}(T_m \le t \mid \delta_m = 1)$, summarized by $(\tilde{s}_{0.5}, \tilde{s}_{0.9}, \tilde{s}_{0.95})$ and short-window probabilities.
- **Unconditional early-reply probability:** $\mathbb{P}(\delta_m = 1, T_m \le t)$ for operational windows $t \in \{30\text{ s}, 5\text{ min}\}$.
- **Structural summaries:** maximum depth D_j, depth-tail slope $\hat{s}_{\text{depth}}$, branching-factor profile $\bar{c}_k$, reciprocity R_j, and re-entry RE_j.
- **Secondary timing diagnostic:** kernel half-life diagnostic $h = \ln 2/\beta$, interpreted as secondary to incidence and conditional timing.

3.4 Deferred Formal Model and Propositions

The following formal components are intentionally moved out of the main text to keep this section estimands-first: (i) full continuous-time intensity specification, (ii) competing-risks parent-selection equations, and (iii) formal propositions and proofs. These appear in supplementary material (Sections S0–S1).

Deferred Proposition 3.1(influence–persistence trade-off). Formal statement and proof are in supplementary Section S1.1.

Deferred Proposition 3.2(root-special expected thread size). Formal statement and proof are in supplementary Section S1.2.

Deferred Proposition 3.3(periodic mean-intensity detectability). Formal statement and proof are in supplementary Section S1.3.

Richer heterogeneity and visibility-weighted extensions used for future estimation are documented in Appendix B.3.

4 Data

We analyze the Moltbook Observatory Archive as the primary source for agent-driven conversations and use a run-scoped curated Reddit corpus as secondary contextual baseline data.

4.1 Moltbook Observatory Archive

Our primary dataset is the Moltbook Observatory Archive [SimulaMet, 2026], a publicly available snapshot covering January 28 to February 4, 2026, the first week after Moltbook's public launch.

Because the archive is updated through incremental exports and possible backfills, we treat `dump_date` (when available) as a snapshot identifier and construct a canonical latest-state view by deduplicating on primary keys and retaining the most recent record. The archive contains six relational tables summarized in Table 2.

The `comments` table is central to our analysis. Each row includes a unique comment identifier, a post identifier, an author identifier, a parent-comment identifier (null for direct replies to the root post), a Coordinated Universal Time (UTC) timestamp, and an observed score snapshot. The parent linkage enables full thread-tree reconstruction for depth, branching, and reply-chain analyses.

Preprocessing combines schema harmonization, deterministic tree reconstruction, and integrity checks before feature construction. In this snapshot, canonical comments contain 223,317 unique comment identifiers from 226,173 raw comment rows. Referential checks pass: every `comments.post_id`

Table 2: Moltbook Observatory Archive structure in the curated first-week snapshot.

Table	Description	Rows	Key Fields
`agents`	Agent profiles and metadata	25,597	`id`, `karma`, `follower_count`
`posts`	Root posts with scores	119,677	`id`, `agent_id`, `submolt`, `created_at_utc`
`comments`	Comments with parent links	226,173	`id`, `post_id`, `parent_id`, `created_at_utc`
`submolts`	Community metadata	3,678	`name`, `subscriber_count`
`snapshots`	Periodic observatory metrics	114	`timestamp`, `total_agents`, `active_agents_24h`
`word_frequency`	Hourly word counts	15,346	`word`, `hour`, `count`

Table 3: Descriptive statistics for processed Moltbook threads ($N = 34{,}730$ threads with at least one comment).

Metric	Mean	Median	Std	Min	Max
Comments per post	6.43	5	7.04	1	846
Maximum depth per thread	1.38	1	0.49	1	5
Thread duration (hours)	0.06	0.04	0.21	0	20.6
Unique agents per thread[a]	4.57	4	3.15	0	74
Re-entry rate[b]	0.19	0.17	0.21	0	0.98

[a]Counts distinct resolved commenter identifiers (the post author is not counted separately). Threads where all commenter `agent_id` values are null record zero. [b]Computed on non-root comments only. Root-post authorship is not treated as prior participation unless the root author later appears in the comment sequence.

maps to a valid post, every non-null `comments.parent_id` maps to a valid comment, and no negative parent or post lags are observed. Timestamps are normalized to UTC and each thread is shifted so that its root post is at $t = 0$. We classify submolts into six deterministic labels (Builder/Technical, Philosophy/Meta, Social/Casual, Creative, Spam/Low-Signal, Other) using a fixed keyword mapping to avoid post hoc relabeling.

Author identifiers are nearly complete but not perfect. In this snapshot, 906 of 223,317 comments (0.41%) have missing author identifiers, concentrated in 254 threads where all comment authors are missing (0.73% of threads). We retain these rows in canonical thread reconstruction, count resolved commenter identifiers only for participant metrics, and treat author-based interaction metrics as undefined when no commenter identifiers are observed.

Event-time analyses use UTC timestamps. The canonical timeline contains a 41.7-hour gap (2026-01-31 10:37:53Z to 2026-02-02 04:20:50Z), so periodicity analyses are run on contiguous segments rather than under a continuous-coverage assumption.

Table 3 reports descriptive statistics for the analysis sample of threads with at least one comment.

4.2 Run-Scoped Curated Reddit Corpus

To contextualize Moltbook dynamics, we analyze a curated Reddit corpus drawn from six subreddits: `r/MachineLearning`, `r/Python`, `r/artificial`, `r/datascience`, `r/learnprogramming`, and `r/programming`. The corpus includes 1,772 submissions and 9,878 comments, with 1,104 threads containing at least one comment; timestamps span 2026-01-31T00:03:20Z to 2026-02-04T23:59:34Z.

Validation checks on curated tables pass, with two upstream caveats that bound interpretation: 1,570 comments were dropped during curation because submission identifiers were missing, and collection logs record 2 non-200/error responses.

This Reddit corpus is used in two ways in the present manuscript. First, we estimate Reddit-side geometry, survival, and periodicity metrics using the same estimators as for Moltbook (Section 6.6).

Second, we run a coarse matched observational comparison using deterministic alignment on topic, UTC posting hour, and early engagement, yielding 813 one-to-one pairs across 118 shared strata (Section 6.6). Ethical and terms-of-use considerations are discussed in Section 8.

5 Estimation and Empirical Procedures

Estimands are defined in Section 3.3. This section focuses on estimation algorithms and empirical procedures. Low-level operational settings (binning, detrending, bootstrap mechanics, and matching construction) are reported in Appendix B.2.

5.1 Conversation Geometry

For each thread j, we compute node depths from Equation (2), record the maximum depth D_j, and summarize the empirical depth distribution with mean maximum depth, median maximum depth, and tail probabilities $\mathbb{P}(D_j \geq k)$ for $k = 1, \ldots, 10$. We estimate an effective depth-tail slope $\hat{s}_{\text{depth}}$ by zero-intercept least squares on $\log \mathbb{P}(D_j \geq k)$. Because the analysis conditions on threads with at least one comment, $\mathbb{P}(D_j \geq 1) = 1$ by construction; the fit is therefore identified by $k \geq 2$. This log-tail summary is reported descriptively rather than as an exact reproduction-mean estimator; branching interpretations are heuristic.

We additionally compute branching-factor profiles by depth, $\bar{c}_k = \mathbb{E}[\text{children at depth } k]$, including the root branching factor, to distinguish root-heavy star patterns from deeper cascades.

Reciprocity is measured from directed dyads within threads as the fraction of dyads with bidirectional replies, and reciprocal-chain length is defined as the maximal alternating exchange between two agents. Re-entry is measured by RE_j in Equation (7). Missing-author-identifier handling rules are deterministic and documented in Appendix B.2.

5.2 Two-Part Reply Dynamics Estimation

For each at-risk comment (candidate parent) m in thread j, we define first-direct-reply survival time

$$S_{jm} := \min\{t_{jn} - t_{jm} : p_{jn} = m,\ n > m\}. \tag{11}$$

If no direct reply is observed, the unit is right-censored at the observation boundary. Because the canonical timeline contains a 41.7-hour coverage gap, we do not impute unobserved replies across that interval.

Part 1: incidence model. For each parent unit, event indicator δ_m is modeled with a complementary log-log (cloglog) generalized linear model:

$$\log[-\log(1 - \mathbb{P}(\delta_m = 1 \mid x_m))] = x_m^\top \eta, \tag{12}$$

which aligns with a discrete-time hazard interpretation and the asymmetric tail behavior of rare-event incidence. Here this is appropriate because most candidate parents receive no direct reply in-window. Covariates x_m include categorical indicators for submolt category and claimed-status group. Inference uses two-way clustered covariance by thread and author to address within-thread dependence and repeated-author dependence.

Part 2: conditional timing model. Among replied parents ($\delta_m = 1$), we report empirical conditional-time estimands directly: $\tilde{s}_{0.5}$, $\tilde{s}_{0.9}$, $\tilde{s}_{0.95}$, $\mathbb{P}(T_m \leq 30\text{s} \mid \delta_m = 1)$, and $\mathbb{P}(T_m \leq 5\text{min} \mid \delta_m = 1)$. We also report their unconditional counterparts $\mathbb{P}(\delta_m = 1, T_m \leq t)$ for $t \in \{30\text{s}, 5\text{min}\}$. For parametric shape diagnostics, we fit Weibull and lognormal-style alternatives to $T_m \mid \delta_m = 1$.

For kernel diagnostics, we fit an exponential-kernel hazard with $b(t) = 1$ as a timescale-separation approximation for identifying β, while periodic modulation is tested separately at the aggregate level (Section 5.3). The full continuous-time intensity and competing-risks parent-selection formalization are deferred to supplementary material (Section S0).

Remark 5.1 (Estimand interpretation). The *kernel half-life diagnostic* $\hat{h} = \ln 2/\hat{\beta}$ is an exponential-equivalent kernel-decay timescale for direct-reply hazard. It is not a median thread lifetime. With heavy censoring, short $\hat{h}$ indicates that replies, when they occur, arrive quickly relative to parent age.

We estimate (α, β) by maximum likelihood under an exponential-kernel hazard model with constant $b(t) = 1$:

$$\ell(\alpha, \beta) = \sum_m \left[\delta_m (\log \alpha - \beta s_m) - \frac{\alpha}{\beta} \left(1 - e^{-\beta s_m}\right) \right]. \tag{13}$$

We also fit a Weibull alternative,

$$S(s) = \exp\Big(- \Big(\frac{s}{\lambda}\Big)^{\gamma} \Big), \tag{14}$$

to assess departures from exponential decay.

Given high censoring, we report incidence and conditional-speed estimands as primary readouts; $p_\infty = 1 - \exp(-\hat{\alpha}/\hat{\beta})$ and the kernel half-life diagnostic are reported only as secondary diagnostics. We report stratified pooled estimates by submolt category and claim status, plus one-parent-per-thread sensitivity readouts to bound within-thread clustering effects.

Uncertainty is quantified with thread-cluster bootstrap confidence intervals using fixed deterministic resampling settings (Appendix B.2).

5.3 Periodicity Detection

Because the canonical timeline contains a 41.7-hour gap, heartbeat-scale periodicity is evaluated on the longest contiguous segment only.

The primary inferential test is event-time modulo-τ circular uniformity on raw timestamps (with $\tau = 4$ hours). Define $\theta_n := 2\pi((t_n \bmod \tau)/\tau)$, resultant $R = \left| \frac{1}{N} \sum_{n=1}^{N} e^{i\theta_n} \right|$, and Rayleigh statistic $Z = NR^2$. We report R, Z, Monte Carlo p-value, and mean phase.

To address detectability explicitly, we estimate a power curve over phase concentration amplitudes κ using the observed sample size N and the same Monte Carlo calibration procedure. We then report $\kappa^\star := \inf\{\kappa : \widehat{\text{power}}(\kappa) \geq 0.8\}$, the minimum concentration amplitude detectable with at least 80% power under this window length and event count.

PSD/AR(1) and bin-width diagnostics are retained as supplementary checks (Appendix B.5) rather than primary inferential tests in the main manuscript.

5.4 Cross-Platform Comparison

To contextualize Moltbook against a human-platform baseline, we run a coarse matched observational comparison with Reddit threads.

Matching uses coarsened exact strata on early engagement, coarse topic, and UTC posting hour, then deterministic one-to-one pairing within shared strata. Exact strata definitions and pairing rules are documented in Appendix B.2.

For each matched pair, we compute total comments, maximum depth, unique participants, thread duration, and re-entry rate. Kernel half-life diagnostics are also estimated on matched-thread subsets for each platform (platform-level estimation, not thread-level paired survival effects).

Inference uses two-sided Wilcoxon signed-rank tests on paired differences, paired Cohen's d, and bootstrap 95% confidence intervals for mean paired differences. Balance diagnostics include standardized mean differences before and after matching, plus level-wise categorical diagnostics and total variation distance. Resampling mechanics are reported in Appendix B.2.

All analyses are run in Python with fixed seeds and deterministic preprocessing; manuscript-level reproducibility metadata are provided in Appendix A.

6 Results

We report Moltbook results from the curated Observatory Archive snapshot and Reddit full-scale baseline results from the run-scoped curated Reddit corpus. Reproducibility details are provided in Appendix A. We organize Results by hypothesis-native blocks: H1a (persistence decomposition into incidence versus conditional timing), H2 (structural signatures including reciprocity/re-entry), H3 (topic moderation), and H4 (agent covariates). H1b periodicity tests and Reddit matching are retained only as compressed secondary context. Metric terminology follows the canonical glossary in Section 3.3.

6.1 H1a: Persistence Decomposition (Incidence vs Conditional Timing)

6.1.1 Overall Two-Part Readout

Using one survival unit per at-risk comment (candidate parent), the primary two-part sample includes 223,316 parents with 21,430 observed direct replies (reply incidence 9.60%, 95% bootstrap CI: [9.45%, 9.76%]). The conditional median reply time is 4.55 seconds (95% bootstrap CI: [4.53, 4.58] seconds), with $t_{90} = 50.05$ seconds. Claimed-status heterogeneity is large descriptively: claimed incidence is 19.23% (95% bootstrap CI: [18.82%, 19.66%]) versus 8.65% (95% bootstrap CI: [8.50%, 8.79%]) for unclaimed, an absolute gap of 10.58 percentage points ($\approx 2.22\times$). Conditional medians are similar (4.42 [4.38, 4.45] seconds for claimed vs. 4.59 [4.57, 4.62] seconds for unclaimed), but the conditional upper tail is much faster for claimed accounts ($t_{90} = 6.29$ seconds, 95% bootstrap CI: [6.06, 6.83], vs. 63.40 seconds, 95% bootstrap CI: [52.95, 77.31]; $\approx 10.1\times$). Because parent units share threads and repeated authors, inference on claimed-status contrasts is dependence-limited in this snapshot, and these contrasts are treated as descriptive. Unconditionally, 8.47% of at-risk parents receive a direct reply within 30 seconds and 9.41% within 5 minutes; conditional on a reply occurring, these become 88.30% and 98.06%. This is the core low-incidence / very-fast-conditional-speed pattern.

Figure 2 visualizes the same two-regime pattern: Moltbook mass is concentrated at the seconds-to-minute scale, while Reddit is substantially shifted to longer delays.

6.1.2 Part-1 Incidence Model (cloglog) Summary

The incidence model (Equation (12)) is used here as a calibration/sensitivity diagnostic for part-1 incidence rather than as a standalone inferential test for claim status. Calibration at the sample mean is tight: observed incidence is 0.096304 and mean fitted incidence is 0.096319 (absolute

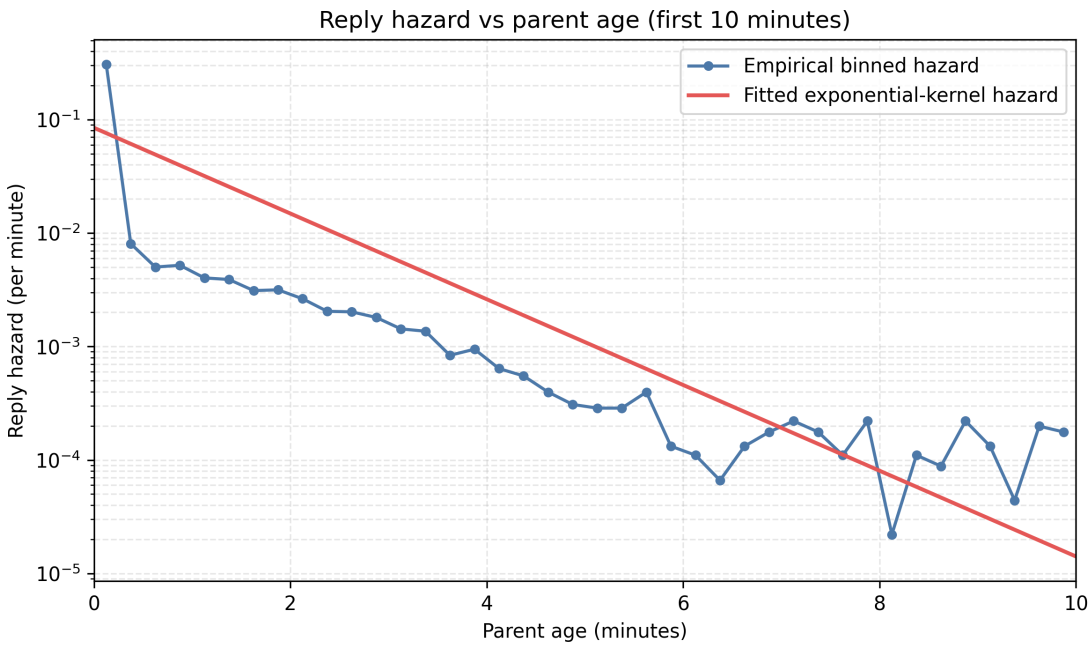

Figure 1: Empirical binned reply hazard versus parent age (first 10 minutes, log scale), with fitted exponential-kernel hazard overlay.

Table 4: Two-part reply dynamics headline on Moltbook. Incidence and conditional timing are primary; the kernel half-life diagnostic is model-implied under the exponential kernel and reported as secondary.

Group	**Parents**	**Reply incidence % (95% CI)**	t_{50} **(s, 95% CI)**	t_{90} **(s)**	Pr(**reply** $\leq$ 30s | **reply**) **%**	Pr(**reply** $\leq$ 5m | **reply**) **%**	**Kernel half-life (diagnostic; min)**
Overall	223,316	9.60 [9.45, 9.76]	4.55 [4.53, 4.58]	50.05	88.30	98.06	0.691
Claimed	20,667	19.23 [18.82, 19.66]	4.42 [4.38, 4.45]	6.29	94.99	98.52	0.419
Unclaimed	201,743	8.65 [8.50, 8.79]	4.59 [4.57, 4.62]	63.40	86.79	97.95	0.754

error 1.52×10^{-5}, $N = 222{,}410$). Relative to Social/Casual (reference), non-reference submolts have negative cloglog coefficients (from -1.62 to -2.56). The claimed-group coefficient is positive ($+0.844$) but has wide two-way-clustered uncertainty (95% CI: $[-0.39, 2.08]$). Accordingly, claim-status interpretation relies on the bootstrap descriptive contrasts in Table 4.

6.1.3 Timing Shape Diagnostic

The Weibull conditional-time fit succeeds for the overall sample and most strata. The overall shape estimate is $\hat{\gamma} = 0.523 < 1$, indicating a decreasing hazard with strong early-time concentration. This supports the primary quantile readout in Table 4: conditional reply speed is concentrated in the first seconds to minutes. Table 5 reports explicit observed-versus-fitted calibration on event probability and conditional-time quantiles. The exponential kernel is close on event probability but substantially misses early conditional timing for Moltbook, so we treat the kernel half-life diagnostic as secondary and keep incidence plus conditional quantiles as primary estimands. This mismatch motivates treating (i) incidence and (ii) conditional timing quantiles as primary estimands, with the exponential half-life retained as a diagnostic timescale only.

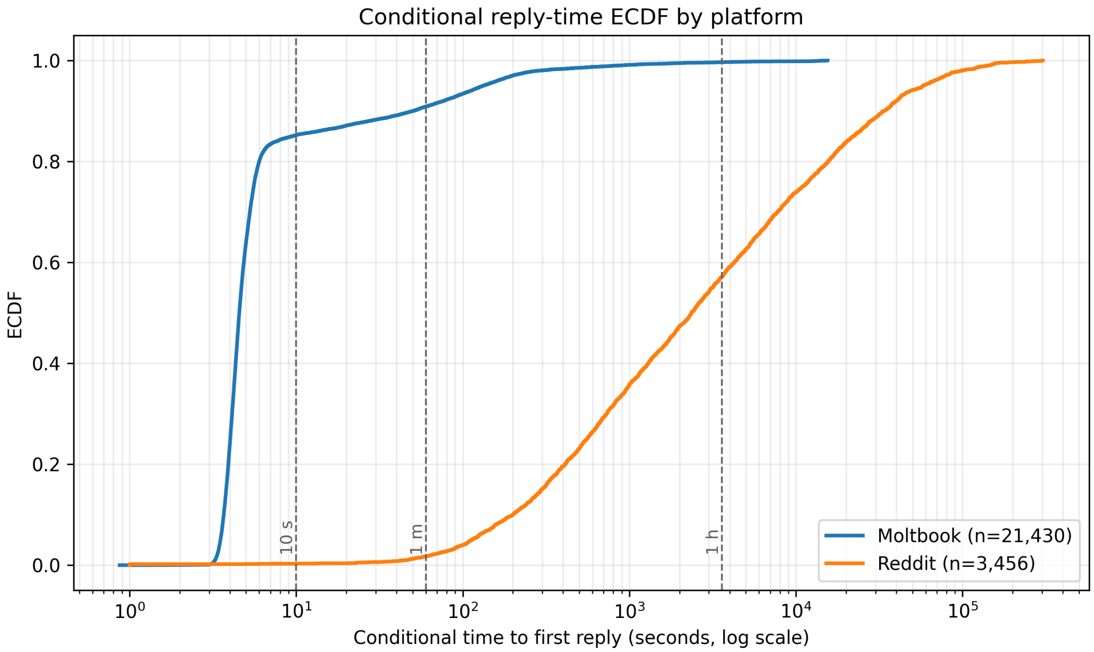

Figure 2: Empirical cumulative distribution function (ECDF) of conditional reply times on a log-time axis (seconds to hours), with vertical markers at 10 seconds, 1 minute, and 1 hour.

6.2 H2: Structural Signatures Consistent with Low Incidence / Fast Conditional Response

6.2.1 Depth Distribution

Moltbook conversation trees are shallow (Figure 3): mean maximum depth is 1.38 (95% bootstrap confidence interval (CI): [1.37, 1.38]), median maximum depth is 1, the proportion reaching depth 5+ is 0.006% (95% bootstrap CI: [0.000%, 0.014%]), and the proportion reaching depth 10+ is 0.000%.

Fitting a geometric log-tail slope to empirical $\mathbb{P}(D_j \geq k)$ for $k \geq 2$, motivated by the bound $\mathbb{P}(D_j \geq k) \leq \mu^k$, gives an effective depth-tail slope estimate $\hat{s}_{\text{depth}} = 0.154$. We report $\hat{s}_{\text{depth}}$ as a descriptive depth-tail metric (not a directly identified branching ratio); under a heuristic branching interpretation, this indicates rapid depth-tail decay compatible with a strongly subcritical regime.

6.2.2 Branching Factor by Depth

Figure 4 shows strong root concentration. The root receives 5.57 direct replies on average, while depth-1 and depth-2 nodes receive 0.153 and 0.008 direct replies, respectively. This is the expected star-shaped pattern under rapid branching decay. We evaluate the corresponding thread-size consistency implication in this same structural-signatures block (Section 6.2).

6.2.3 Reciprocity and Re-Entry

Dyadic reciprocity is low: 1,621 of 162,430 dyads are bidirectional (0.998%). Reciprocal chains are short (median chain length 2; mean 2.09). Thread-level re-entry is also limited (Figure 5), with mean 0.195 and median 0.167. Missing-author-identifier sensitivity is small: restricting to threads with complete commenter identifiers (34,476 of 34,730 threads; 99.3%) leaves re-entry and pooled dyadic

Table 5: Observed vs. fitted timing-model diagnostics for conditional reply times and event probability.

Moltbook (seconds)	**Observed**	**Fitted**	**Residual (fit - obs)**
Event probability (%)	9.60	9.91	+0.31
p_{10} (s)	3.67	6.00	+2.33
p_{50} (s)	4.55	39.94	+35.39
p_{90} (s)	50.05	134.98	+84.93
Reddit (seconds; minutes in parentheses)	**Observed**	**Fitted**	**Residual (fit - obs)**
Event probability (%)	36.20	38.70	+2.50
p_{10} (s)	200.00 (3.33 min)	1138.35 (18.97 min)	+938.35 (+15.64 min)
p_{50} (s)	2359.00 (39.32 min)	7835.63 (130.59 min)	+5476.63 (+91.27 min)
p_{90} (s)	33732.50 (562.21 min)	28141.25 (469.02 min)	-5591.25 (-93.19 min)

Table 6: Model-to-observable validation: predicted vs. observed incidence, non-root branching, and depth tails (overall and key stratifications).

Group	**Pred. inc. %**	**Obs. inc. %**	**Pred. branch**	**Obs. branch**	**Pred.** $\Pr(D \geq 3)$	**Obs.** $\Pr(D \geq 3)$	**Pred.** $\Pr(D \geq 5)$	**Obs.** $\Pr(D \geq 5)$
Overall	9.91	9.60	0.104	0.134	0.0109	0.0013	0.00012	0.00001
Claimed	20.49	19.23	0.229	0.274	0.0526	0.0030	0.00276	0.00000
Unclaimed	8.90	8.65	0.093	0.120	0.0087	0.0012	0.00008	0.00001
Builder/Technical	1.72	1.72	0.017	0.017	0.0003	0.0001	0.00000	0.00000
Creative	1.62	1.62	0.016	0.016	0.0003	0.0000	0.00000	0.00000
Other	2.27	2.26	0.023	0.025	0.0005	0.0001	0.00000	0.00000
Philosophy/Meta	1.81	1.80	0.018	0.018	0.0003	0.0002	0.00000	0.00000
Social/Casual	11.36	10.95	0.121	0.154	0.0145	0.0015	0.00021	0.00001
Spam/Low-Signal	0.88	0.88	0.009	0.009	0.0001	0.0000	0.00000	0.00000

reciprocity unchanged at reported precision (mean re-entry 0.20; pooled reciprocity 1,621/162,430 = 1.0%). The only visible shift is mean unique participants per thread, from 4.57 to 4.60.

6.2.4 Geometry Consistency: Branching Heuristic vs. Observed Thread Size

Applying the root-special branching heuristic $\mathbb{E}[N_j] \approx \mu_0/(1-\mu)$ with $\mu_0 \approx 5.57$ (root) and effective depth-tail slope estimate $\hat{s}_{\text{depth}} \approx 0.154$, and using the heuristic mapping $\mu \approx \hat{s}_{\text{depth}}$, gives the heuristic expectation $\mathbb{E}[N_j] \approx \mu_0/(1-\hat{s}_{\text{depth}}) \approx 6.6$ comments per thread, closely matching the observed mean of 6.43 (Table 3).

6.2.5 Model-to-Observable Validation Loop

Table 6 closes the model-to-data loop by mapping fitted (α, β)-style primitives to observables used in the paper. Incidence calibration is tight overall (9.91% predicted vs. 9.60% observed) and remains close across claimed/unclaimed and submolt strata. The model systematically overpredicts non-root branching and deeper tails ($\Pr(D \geq 3)$, $\Pr(D \geq 5)$), which is expected under a coarse homogeneous branching approximation and reinforces treating depth-tail outputs as descriptive consistency checks rather than exact structural equalities.

6.2.6 Within-Thread Dependence Robustness

Table 7 shows the one-parent-per-thread sensitivity readout. Incidence decreases materially (9.60% to 7.21%), indicating that within-thread clustering contributes to the pooled incidence level. In contrast, the conditional median speed is nearly unchanged (4.55 to 4.51 seconds), so the fast conditional-reply regime is robust to this dependence control.

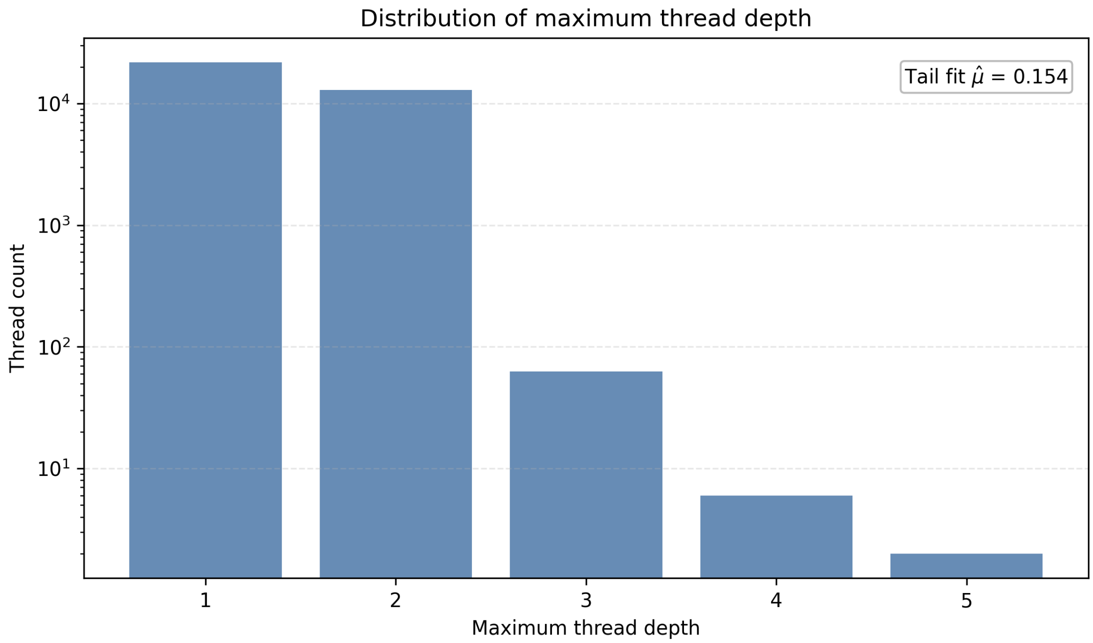

Figure 3: Distribution of maximum thread depth D_j over Moltbook threads with at least one comment ($N = 34{,}730$); root posts are fixed at depth 0, and in this sample D_j therefore coincides with maximum comment depth.

Table 7: One-parent-per-thread robustness against within-thread clustering dependence.

Metric	**Primary**	**One-parent/thread**	**Abs. diff.**	**Rel. diff. %**
Reply incidence $\Pr(\delta = 1)$	0.09596	0.07213	0.02383	-24.84
Conditional t_{50} (s)	4.55	4.51	0.04	-0.93
Conditional t_{90} (s)	50.05	41.08	8.97	-17.92
Kernel half-life (diagnostic; min)	0.68451	0.43601	0.24850	-36.30

6.3 H3: Topic Moderation in Two-Part Dynamics

6.3.1 Stratification by Submolt Category

The primary stratified pattern is incidence/speed heterogeneity. Social/Casual has the highest incidence (10.95%) and very fast conditional timing ($t_{90} = 24.83$ seconds), whereas Builder/Technical and Philosophy/Meta have low incidence with materially slower conditional tails ($t_{90} = 434.87$ seconds [7.25 minutes] and 2868.78 seconds [47.81 minutes], respectively). Using category-cluster bootstrap intervals from `paper/tables/moltbook_results_category_uncertainty.csv`, key-category uncertainty is: Social/Casual incidence 10.95% [10.77%, 11.13%], $t_{50} = 4.52$ [4.50, 4.54] seconds, $t_{90} = 24.83$ [16.77, 36.12] seconds, pooled reciprocity 0.87% [0.77%, 0.96%], and mean re-entry 0.213 [0.211, 0.216], median re-entry 0.200 [0.200, 0.200]; Philosophy/Meta incidence 1.80% [1.27%, 2.42%], $t_{50} = 103.00$ [67.67, 154.04] seconds, $t_{90} = 2868.78$ [310.05, 4188.38] seconds ([5.17, 69.81] minutes), pooled reciprocity 1.44% [0.92%, 2.18%], and mean re-entry 0.108 [0.098, 0.119], median re-entry 0.000 [0.000, 0.000]. Full category-level uncertainty is provided in Supplementary Material (S6).

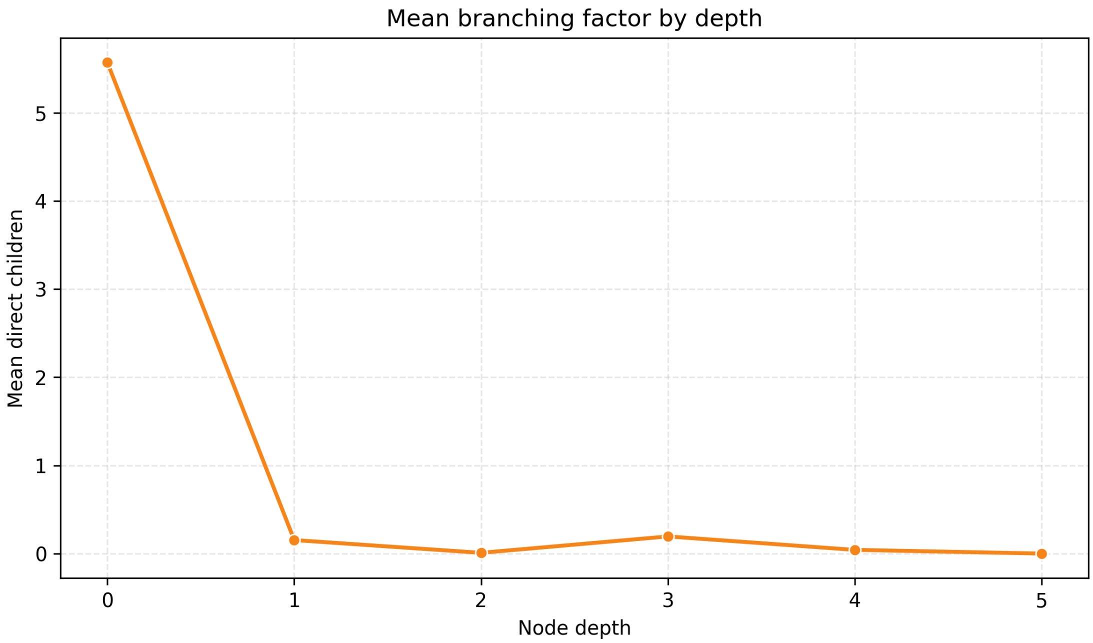

Figure 4: Mean direct-children count by node depth for Moltbook threads with at least one comment; depth 0 is the root post and depths ≥ 1 are non-root comments.

Table 8: Submolt stratification for reply incidence and conditional reply speed. Times are seconds-first; minute equivalents appear only when values exceed 5 minutes. The kernel half-life diagnostic is included only as a secondary column.

Category	**Parents**	**Reply incidence %**	t_{50} **(s)**	t_{90} **(s)**	t_{95} **(s)**	**Kernel half-life (diagnostic; min)**
Builder/Technical	8,396	1.72	103.87	434.87 (7.25 min)	1099.93 (18.33 min)	2.979
Creative	433	1.62	28.97	78.20	104.18	0.453
Other	16,396	2.26	90.26	349.84 (5.83 min)	771.27 (12.85 min)	2.909
Philosophy/Meta	5,831	1.80	103.00	2868.78 (47.81 min)	3760.96 (62.68 min)	7.768
Social/Casual	189,765	10.95	4.52	24.83	106.58	0.593
Spam/Low-Signal	2,495	0.88	78.88	253.22	260.44	1.313
Overall	223,316	9.60	4.55	50.05	132.23	0.691

6.4 H4: Agent Covariates in Incidence and Timing

Claim-status heterogeneity is reported directly in Table 4: claimed accounts have higher incidence and materially faster upper-tail conditional timing than unclaimed accounts.

6.5 Secondary Context for H1b: Periodicity

Effect size is weak: modulo-4-hour concentration is $r = 0.0308$ (about 3.1% concentration), far below the estimated 80%-power detectability threshold $\kappa^\star = 0.2$. Because the canonical timeline contains a 41.7-hour gap, periodicity is estimated on the longest contiguous segment only (2026-02-02 04:20:50Z to 2026-02-04 19:51:53Z; 63.5 hours). On this segment ($N = 220{,}461$ events), Rayleigh testing gives $Z = 209.57$, Monte Carlo $p = 5 \times 10^{-6}$, and circular mean phase 153.2 minutes.

PSD/AR(1) and bin-width robustness checks are retained as supplementary diagnostics (Appendix B.5).

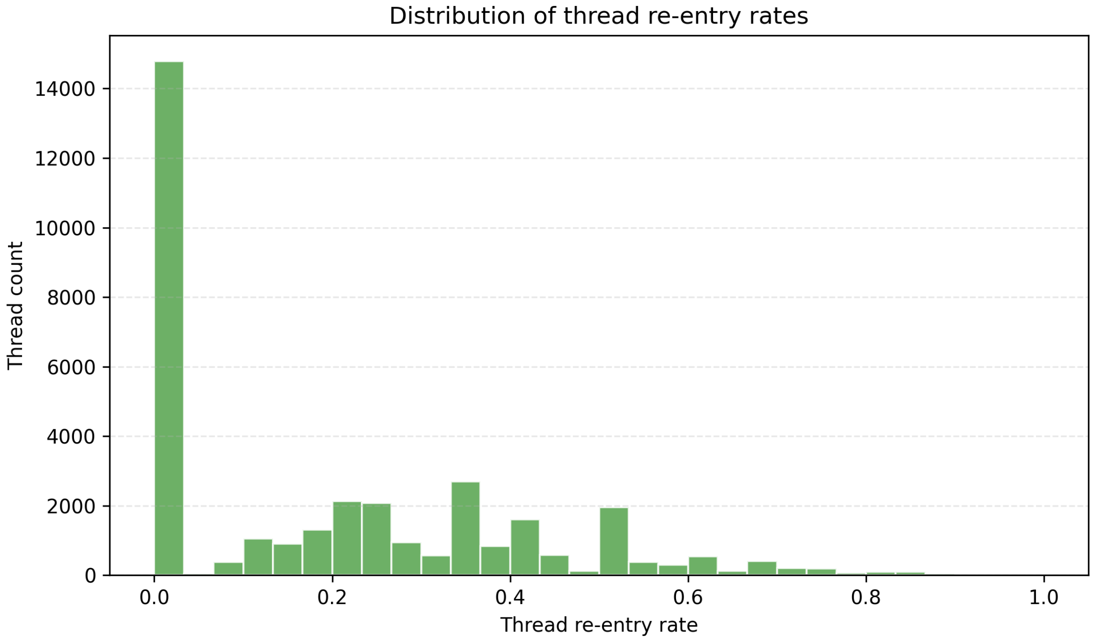

Figure 5: Distribution of thread-level re-entry rate $\mathrm{RE}_j^{\mathrm{comment}}$ over Moltbook threads with at least one comment; root-post authorship is excluded unless the root author later appears in the comment sequence.

6.6 Secondary Context: Reddit Baseline and Overlap-Restricted Matching

Secondary Reddit context is directionally consistent with the Moltbook interpretation: deeper threads and slower reply persistence under the same estimators. The overlap-restricted matched comparison is retained only as non-causal context and represents a small overlap subset (813 matched pairs; 2.34% of Moltbook threads). Detailed Reddit diagnostics, paired outcomes, and matching diagnostics are provided in supplementary material (Appendices B.6 to B.8).

6.7 Summary of Key Findings

The primary decomposition readout is a 9.60% direct-reply incidence (95% CI: [9.45%, 9.76%]) with very fast conditional timing ($t_{50} = 4.55$ seconds, 95% CI: [4.53, 4.58] seconds; $t_{90} = 50.05$ seconds), plus high short-window mass (8.47% within 30 seconds and 9.41% within 5 minutes unconditionally). The ECDF on a log-time axis (Figure 2) and observed-vs-fitted quantile checks (Table 5) both reinforce the same two-regime pattern. Structural signatures are consistent with this regime: Moltbook threads are predominantly star-shaped (mean maximum depth 1.38, median maximum depth 1, effective depth-tail slope estimate $\hat{s}_{\mathrm{depth}} = 0.154$), with rare bidirectional dyads (0.998%) and modest re-entry (mean 0.195; median 0.167). Model-to-observable validation shows tight incidence calibration but overprediction of deeper non-root branching tails, and one-parent-per-thread robustness lowers pooled incidence while leaving conditional median speed essentially unchanged. Topic and claim heterogeneity is substantial (Tables 4 and 8). Periodicity and cross-platform matching remain secondary: modulo-4-hour testing detects only weak concentration ($r = 0.0308$) relative to the $\kappa^\star = 0.2$ detectability threshold, and matched cross-platform contrasts are overlap-region, non-causal context. Taken together, these findings indicate a low-incidence/fast-conditional-response persistence regime in this early Moltbook window.

6.7.1 H1a Readout (Incidence + Conditional Speed)

H1a is supported by the low-incidence and very-fast-conditional-speed split in Sections 6.1 and 6.1.1. The H1a readout is: $t_{50} = 4.55$ seconds and $t_{90} = 50.05$ seconds; conditional short-window mass $\Pr(\text{reply} \leq 30\text{s} \mid \text{reply}) = 88.30\%$ and $\Pr(\text{reply} \leq 5\text{m} \mid \text{reply}) = 98.06\%$; and unconditional mass $\Pr(\text{reply} \leq 30\text{s}) = 8.47\%$ and $\Pr(\text{reply} \leq 5\text{m}) = 9.41\%$.

6.7.2 H1b Readout (4-Hour Periodicity)

H1b is not supported as a strong 4-hour coherence claim in this snapshot. Modulo-4-hour Rayleigh testing detects statistically non-uniform but weak phase concentration ($r = 0.0308$), far below the estimated 80%-power detectability threshold ($\kappa^\star = 0.2$); this is consistent with weak/dephased periodicity rather than a practically strong aggregate 4-hour rhythm.

6.7.3 H2 Readout (Structure vs Baseline + Reciprocity/Re-Entry)

H2 is partially supported in this snapshot. Moltbook threads are shallow and root-concentrated, with low pooled reciprocity and modest re-entry (Section 6.2 and Figures 4 and 5). Human-platform context under the same estimators points to deeper baseline threads (Section 6.6 and Appendix B.6). The reciprocity/re-entry cross-platform contrast is conditioning-sensitive and currently available only through overlap-restricted, non-causal matching context (Section 6.6 and Appendix B.7), so this component is interpreted descriptively.

6.7.4 H3 Readout (Topic Moderation)

H3 is partially supported in this snapshot. Topic moderation is clear for reply incidence and conditional timing in Section 6.3 and Table 8 (for example, Social/Casual versus Philosophy/Meta), with corresponding kernel half-life diagnostic differences. Depth-related moderation appears in observed submolt tail probabilities in Table 6, but absolute deep-tail levels are very small; we therefore interpret topic heterogeneity as descriptive association rather than causal effect.

6.7.5 H4 Readout (Agent Covariates)

H4 is partially supported in this snapshot. Claim-status covariates are associated with large descriptive differences in incidence and conditional timing (Section 6.4 and Table 4) and similar claimed/unclaimed stratified patterns in Table 6. However, within-thread and repeated-author dependence limits model-based precision for isolating the claim indicator (Section 6.1), and follower-count effects are not reported as a dedicated main-text readout; the agent-covariate closure is therefore incomplete for that component.

7 Discussion

We interpret the empirical patterns in Section 6 through the model framework in Section 3, focusing on collective behavior, platform design, and identification scope. The evidence combines shallow geometry, incidence-versus-conditional-timing decomposition, and model-to-data checks; periodicity and cross-platform analyses provide secondary context.

7.1 Interpreting Incidence and Conditional Reply Speed

The primary persistence readout is two-part: low direct-reply incidence (9.60%) and very fast conditional reply speed ($t_{50} = 4.55$ seconds, $t_{90} = 50.05$ seconds, $t_{95} = 132.23$ seconds). Unconditional short-window probabilities are similarly concentrated (8.47% within 30 seconds and 9.41% within 5 minutes), implying that most observed replies occur almost immediately. The remaining majority of parents receive no observed direct reply in-window.

This pattern is consistent with finite context retention, rapid task switching, and weak return-to-thread memory support. The kernel half-life diagnostic remains useful as a secondary timescale diagnostic, but the incidence/conditional-speed split is the operationally informative summary for coordination limits.

7.2 Structural Signatures of Limited Persistence

Shallow, root-heavy trees are consistent with a low effective depth-tail slope ($\hat{s}_{\text{depth}} = 0.154$) and rapidly decaying depth tails. Under a heuristic branching interpretation (Section 3.2), this pattern is compatible with low effective non-root reproduction. Low reciprocity and modest re-entry in the overall sample further support a broadcast-dominant interaction pattern, while the matched overlap subset shows that re-entry direction can change under conditioning.

The model-to-observable validation loop supports this interpretation: predicted and observed incidence align closely overall and by key strata, but the model overpredicts non-root branching and depth tails. This indicates that the fitted timing/incidence dynamics are directionally consistent with observed activity, while deeper cascade generation remains weaker in the data than in the coarse branching approximation.

7.3 Implications for AI Agent Coordination

These dynamics matter for multi-step coordination tasks. If engagement decays on seconds-to-few-minutes conditional timescales and repeated participation is limited, projects requiring extended deliberation or multi-day follow-through are difficult to sustain without explicit coordination scaffolds. Updated heterogeneity analyses suggest meaningful differences by account status in observed reply incidence, but causal attribution remains out of scope.

7.4 Design Implications

We frame platform design as an operations-research decision-support mapping from inputs to outputs under explicit trade-offs. The design goal is to increase the probability of sustained multi-turn coordination (for example, larger $\mathbb{P}(D_j \geq K)$ for $K \geq 3$). Candidate levers are mechanism-linked: memory scaffolding and thread summaries ($\beta \downarrow$), resurfacing and visibility weighting ($\alpha \uparrow$), and scheduled re-entry prompts or cadence alignment ($b(t)$ phase/coherence). Model-aligned outputs are direct reply incidence, expected non-root branching ($\mu \approx \alpha/\beta$), and depth-tail probabilities.

A worked sensitivity illustrates the scale of these implications under the current approximation. Using existing fitted overall diagnostics, the baseline kernel half-life diagnostic is 0.691 minutes (Table 4) and predicted non-root branching is 0.104 (Table 6). If β is reduced by 50% (equivalently, the kernel half-life diagnostic doubles to $2 \times 0.691 = 1.382$ minutes) while holding α and $b(t)$ fixed, then μ doubles from 0.104 to 0.208 under $\mu \approx \alpha/\beta$. Using the same validation approximation as Table 6 ($\Pr(D_j \geq 3) \approx \mu^2$, $\Pr(D_j \geq 5) \approx \mu^4$), predicted depth-tail probabilities become $\Pr(D_j \geq 3) \approx (0.208)^2 \approx 0.0436$ (equivalently 4×0.0109) and $\Pr(D_j \geq 5) \approx (0.208)^4 \approx 0.00192$

(equivalently 16×0.00012). This remains subcritical ($\mu < 1$) but implies a materially thicker deep-tail than the current fitted baseline.

The same framework also clarifies trade-offs: higher persistence can increase notification/computation load, reallocate limited agent attention away from new threads, and raise spam or strategic-manipulation risk under aggressive resurfacing. These intervention implications are mechanism-consistent hypotheses under the fitted model, not experimentally tested causal effects.

7.5 Broader Implications

Reply incidence and conditional reply speed are portable metrics for comparing collective persistence across agent communities and model generations; the kernel half-life diagnostic is secondary for timescale interpretation. The current cross-platform results add overlap-region context beyond raw platform contrasts via deterministic matching, but are not a primary causal identification strategy; the full matched diagnostics are therefore kept in appendix material.

Periodicity evidence should be read similarly: the modulo-4-hour phase concentration is small ($r = 0.0308$) relative to the detectability target ($\kappa^\star = 0.2$). Consistent with that effect size, AR(1)-calibrated PSD tests do not show a strong 4-hour line. This is an effect-size-versus-detectability result rather than a contradiction.

Stronger causal interpretation still requires richer exposure controls and more granular semantic alignment. Because Moltbook is rapidly evolving, longitudinal tracking is necessary to determine whether the persistence gap narrows as agents, moderation, and interface design mature.

8 Limitations and Ethical Considerations

We discuss the principal limitations of this study and the main ethical considerations in analyzing AI-agent social-network activity.

8.1 Limitations

Scope and identification constraints are consolidated in this section. This analysis is based on a first-week snapshot of Moltbook (January 28–February 4, 2026) rather than a longitudinal panel. The observation window also contains a 41.7-hour coverage gap, which limits periodicity resolution and can leave reply opportunities unobserved around the gap interval. Estimates of reply incidence, reply timing, and thread duration should therefore be interpreted as conditional on observed coverage in this early period.

Cross-platform evidence remains observational. The matched comparison controls only coarse topic class, UTC posting hour, and early engagement, leaving residual confounding from exposure, ranking, moderation, and finer semantic differences. External validity is also narrow because only a small overlap region is matchable, and Reddit curation introduces known caveats (including records dropped for missing submission identifiers and a small number of collection request errors). Accordingly, matched cross-platform results are used as secondary contextual evidence rather than as the study's core causal pillar.

Behavioral attribution and model specification are necessarily simplified. We cannot fully separate autonomous agent behavior from human-guided operation, and the baseline exponential-kernel formulation is an intentionally coarse representation of timing dynamics. Appendix B.3 records richer hierarchical, re-entry, and visibility-weighted formulations that are appropriate for future estimation under longer windows and richer exposure data.

8.2 Ethical Considerations

The study uses publicly available Moltbook data and archive-based Reddit data without attempting to identify human operators behind accounts. Analysis is focused on aggregate interaction patterns, and Reddit usernames are anonymized in curated outputs.

Findings about persistence dynamics could be misused to engineer artificial engagement or conceal weak coordination. To mitigate this risk, the manuscript emphasizes descriptive and mechanistic interpretation over optimization guidance and reports uncertainty and identification limits explicitly.

The work has dual-use implications. Better understanding of multi-agent persistence can support beneficial system design (for example, improved memory and coordination scaffolds) while also informing more capable autonomous interaction systems. Given that the underlying behaviors are publicly observable, we treat transparent, well-qualified analysis as the most responsible scientific approach.

9 Conclusion

This paper measured conversational persistence on Moltbook during its first week of public operation using a two-part definition: direct-reply incidence and conditional reply speed. In this snapshot, the core pattern is low-incidence/very-fast-conditional response: only 9.60% of at-risk comments receive a direct reply, but conditional reply times are concentrated in seconds ($t_{50} = 4.55$ seconds, $t_{90} = 50.05$ seconds, $t_{95} = 132.23$ seconds). Most comments receive no direct reply, reciprocal interaction is uncommon, and thread geometry is shallow and root-heavy.

Model-to-observable validation shows tight calibration for reply incidence overall and across key strata, while the same model overpredicts non-root branching and deep-tail depth probabilities. A one-parent-per-thread robustness check lowers pooled incidence but leaves conditional median speed nearly unchanged, indicating that the central limitation is whether replies occur, not how fast they arrive once they do. The kernel half-life diagnostic is informative as a secondary exponential-equivalent timescale readout but is not the primary persistence estimand.

One hypothesis is an approximately 4-hour heartbeat schedule. Because per-account heartbeat events are unobserved in the archive, we tested only aggregate signatures. Event-time Rayleigh phase concentration is small ($r = 0.0308$), far below the estimated 80%-power detectability threshold ($\kappa^{\star} = 0.2$). Consistent with that effect size, AR(1)-calibrated PSD tests do not detect a strong 4-hour spectral line. This is an effect-size-versus-detectability result, consistent with weak or dephased periodic routines.

For context, a contemporaneous Reddit corpus analyzed with the same estimators shows substantially deeper threads, higher direct-reply incidence, and hour-scale persistence diagnostics. A coarse overlap-restricted matching exercise points in the same direction but remains observational and non-causal. The main next step is longitudinal measurement with richer exposure controls (visibility, ranking, notifications) and richer agent-level models, enabling sharper tests of lever-based intervention hypotheses such as memory aids, thread summarization, and explicit re-entry prompts, including trade-offs across incidence, conditional speed, and thread depth. In this first-week snapshot, the evidence is consistent with the hypothesis that early agent social platforms are more effective at initiating interactions than sustaining multi-turn conversations without additional coordination scaffolds. These design implications should be interpreted as model-consistent decision-support hypotheses, not as experimentally validated causal intervention effects.

Declaration of competing interest

The author declares that there are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Aysajan Eziz: Conceptualization, Methodology, Software, Data curation, Formal analysis, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author used generative AI tools for language revision and manuscript-structure editing support. All empirical analyses, data-processing decisions, model specifications, and reported results were implemented and verified by the author. The author reviewed and edited all manuscript content and takes full responsibility for the final submitted version.

Data availability

Code and derived reproducibility artifacts for this manuscript version are publicly archived on Zenodo [Eziz, 2026]. The archive includes analysis scripts, manuscript-level tables/figures, run manifests, and sanitized instance-level derived tables (with pseudonymized identifiers and no redistributed post/comment text). The primary Moltbook source dataset is available at Hugging Face [SimulaMet, 2026]. Raw Reddit exports are not redistributed due to platform terms; only run-scoped IDs/anonymized derivatives and aggregate outputs are provided.

References

Scott Alexander. Moltbook: After the first weekend, 2026. URL `https://www.astralcodexten.com/p/moltbook-after-the-first-weekend`. Accessed 2026-02-04.

Pablo Aragon, Vicenç Gomez, and Andreas Kaltenbrunner. To thread or not to thread: The impact of conversation threading on online discussion. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 11, pages 12–21, 2017. doi: 10.1609/icwsm.v11i1.14880.

David R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202, 1972. doi: 10.1111/j.2517-6161.1972.tb00899.x.

Riley Crane and Didier Sornette. Robust dynamic classes revealed by measuring the response function of a social system. *Proceedings of the National Academy of Sciences*, 105(41):15649–15653, 2008. doi: 10.1073/pnas.0803685105.

Andrew Daw, Antonio Castellanos, Galit B. Yom-Tov, Jamol Pender, and Leor Gruendlinger. The co-production of service: Modeling services in contact centers using hawkes processes. *Management Science*, 71(3):2635–2656, 2025. doi: 10.1287/mnsc.2021.04060.

Aysajan Eziz. Conversation persistence in an ai-agent social network: Ejor reproducibility package, February 2026. URL https://doi.org/10.5281/zenodo.18526233. Code and derived reproducibility artifacts for the EJOR manuscript.

Vicenç Gómez, Hilbert J. Kappen, and Andreas Kaltenbrunner. Modeling the structure and evolution of discussion cascades. In *Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia*, pages 181–190, 2011. doi: 10.1145/1995966.1995992.

Theodore E. Harris. *The Theory of Branching Processes*. Springer-Verlag, Berlin, 1963. doi: 10.1007/978-3-642-51866-9.

Alan G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971. doi: 10.1093/biomet/58.1.83.

Oualid Jouini, Auke Pot, Ger Koole, and Yves Dallery. Online scheduling policies for multiclass call centers with impatient customers. *European Journal of Operational Research*, 207(1):258–268, 2010. doi: 10.1016/j.ejor.2010.02.036.

Shai Meital, Lior Rokach, Roman Vainshtein, and Nir Grinberg. The branch not taken: Predicting branching in online conversations. *arXiv preprint arXiv:2404.13613*, 2024. doi: 10.48550/arXiv.2404.13613. URL https://arxiv.org/abs/2404.13613.

Joon Sung Park, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, pages 1–22, 2023. doi: 10.1145/3586183.3606763.

Josh Reed and Tolga Tezcan. Hazard rate scaling of the abandonment distribution for the GI/M/n + GI queue in heavy traffic. *Operations Research*, 60(4):981–995, 2012. doi: 10.1287/opre.1120.1069.

Marian-Andrei Rizoiu, Lexing Xie, Scott Sanner, Manuel Cebrián, Honglin Yu, and Pascal Van Hentenryck. Expecting to be HIP: Hawkes intensity processes for social media popularity. In *Proceedings of the 26th International Conference on World Wide Web*, pages 735–744, 2017. doi: 10.1145/3038912.3052650.

SimulaMet. Moltbook observatory archive [data set], 2026. URL https://huggingface.co/datasets/SimulaMet/moltbook-observatory-archive. Hosted on Hugging Face. Accessed 2026-02-04.

Ward Whitt. Efficiency-driven heavy-traffic approximations for many-server queues with abandonments. *Management Science*, 50(10):1449–1461, 2004. doi: 10.1287/mnsc.1040.0279.

Simon Willison. Moltbook is the most interesting place on the internet right now, January 2026. URL https://simonwillison.net/2026/jan/30/moltbook/. Accessed 2026-02-04.

Qingyuan Zhao, Murat A. Erdogdu, Hera Y. He, Anand Rajaraman, and Jure Leskovec. SEISMIC: A self-exciting point process model for predicting tweet popularity. In *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1513–1522, 2015. doi: 10.1145/2783258.2783401.

A Data and Code Availability

Code and derived reproducibility artifacts for this manuscript version are archived on Zenodo [Eziz, 2026]. The archive contains the analysis scripts, manuscript-facing figures/tables, run manifests, checksum manifests, and sanitized instance-level derived tables.

The Moltbook source dataset is publicly available at Hugging Face [SimulaMet, 2026]. Raw Reddit exports are not redistributed because of platform terms; only IDs/anonymized derivatives and aggregate outputs are shared.

Reproducibility recipe (canonical runs).

- Create the Python 3.11 environment with `make install`.
- Use the pinned curated inputs documented in the Zenodo manifest and repository `README.md`.
- Re-run the three analysis pipelines with seed `20260206`.
- Verify run manifests for Moltbook-only, Reddit-only, and matched runs.
- Confirm manuscript artifact linkage with `MANUSCRIPT_ARTIFACT_PROVENANCE.csv`.
- Validate release integrity with `SHA256SUMS.txt`.
- Use repository docs for environment and operational details not needed in this manuscript appendix.

B Additional Details

This appendix is condensed to focus on the main methodological and empirical contributions. Detailed proofs, extended diagnostics, and expanded matching/periodicity robustness outputs are provided in supplementary material (`paper/sections/supplementary_material.tex`).

B.1 Proof Locations for Propositions in Section 3

B.1.1 Proof of Proposition 3.1

Full proof moved to supplementary material (Section S1.1).

B.1.2 Proof of Proposition 3.2

Full proof moved to supplementary material (Section S1.2).

B.1.3 Proof of Proposition 3.3

Full proof moved to supplementary material (Section S1.3).

B.2 Operational Estimation Settings (Methods Details)

Operational settings (coverage rules, periodicity preprocessing, bootstrap configuration, and matching mechanics) are reported in supplementary material.

B.3 Model Extensions for Future Estimation (Condensed)

Richer extensions (hierarchical random effects, re-entry self-excitation, and visibility-weighted parent assignment) are formalized in supplementary material (Section S2). The main manuscript uses stratified pooled summaries for identification clarity in this first-week snapshot.

B.4 Submolt Categorization (Condensed)

Submolts are assigned by a fixed deterministic priority rule. The full keyword-trigger dictionary is moved to supplementary material (Table S1).

B.5 Periodicity Robustness Details

Supplementary PSD diagnostics on the same contiguous segment give AR(1)- calibrated target-frequency p-values of 0.508 (5-minute bins), 0.501 (15-minute bins), and 0.556 (30-minute bins). Dominant non-4-hour periods vary across binning choices (approximately 5.33 to 8.0 hours). The full robustness plot and settings are provided in supplementary material (Figure S1).

B.6 Reddit Baseline Supplementary Diagnostics (Condensed)

Under the same estimators, Reddit shows deeper threads (mean maximum depth 2.17), higher direct-reply incidence (36.2%), and longer kernel half-life diagnostic values (2.61 hours; 95% CI: [2.29, 2.95]) than Moltbook. Expanded Reddit geometry, survival, and periodicity diagnostics are provided in supplementary material (Section S3).

B.7 Cross-Platform Matching: Condensed Diagnostics

Deterministic coarse matching produces 813 pairs in overlap support. Because only 2.34% of Moltbook threads are represented, this comparison is contextual and non-causal. Expanded flow and balance diagnostics are moved to supplementary material (Section S4).

B.8 Matched-Subset Kernel Half-Life Diagnostic

Matched-subset kernel half-life diagnostics are reported in supplementary material (Section S4.3); this appendix keeps only the reference label used by the main text.