Datamart-Agent: LLM-Driven Game-Theoretic Agent for Data Marketplace Modeling

Pangjing Wu[1], Peter-Q. Chen[1], Xiaodong Li[2], Wenqi Fan[1], Qing Li[1]*,
[1]The Hong Kong Polytechnic University, [2]Hohai University

Abstract

Data marketplaces analyze strategic data exchange among users, platforms, and buyers. However, most existing studies model on static equilibria and complete information, which limits their realism. In this work, we study whether large language model (LLM)-driven agents can make equilibrium-consistent decisions in analytically tractable data marketplaces with evolving and incomplete-information. Specifically, we introduce **EvoDM**, an agent-based modeling framework that extends the classical static data marketplace to dynamic and incomplete-information settings while providing tractable equilibrium benchmarks for evaluating agent decisions. Building upon EvoDM, we propose **Datamart-Agent**, an LLM-driven game-theoretic agent that improves equilibrium-consistent decision execution through dynamic game tree memory and mechanism-guided reflection, without requiring parameter updates. Experiments demonstrate that Datamart-Agent closely matches equilibrium-consistent decision-making, achieving the lowest utility gap and over 20% higher Pass@ϵ than strong baselines. After validating its effectiveness, we employ EvoDM with Datamart-Agent to analyze competition and regulation in assumption-relaxed settings where closed-form ground truth is unavailable, providing exploratory simulation-based insights into market dynamics and regulatory effects.

1 Introduction

Data marketplaces play a central role in modern AI infrastructure by enabling data sharing and reuse. As large language models (LLMs) scale and publicly available datasets approach saturation (Aghajanyan et al., 2023), high-value proprietary data has become scarce (Chen et al., 2025a). This scarcity creates monetary incentives for organizations to collect and share higher-quality data (Agarwal et al., 2019; Zhu et al., 2024), while buyers turn to purchase these data products from marketplaces for competitive advantage (Borgman, 2012). Consequently, three-layer data marketplaces, such as AWS Data Exchange (aws, 2025), have attracted growing attention from academia and industry (Moor, 2019; Fernandez et al., 2020a; Fallah et al., 2024).

Existing studies on data marketplaces have primarily focused on formalizing data trading mechanisms and incentive structures (Castro Fernandez, 2023), with later work incorporating privacy considerations driven by regulations such as the GDPR (Huang et al., 2023; Xiao et al., 2023). Building on these foundations, equilibrium analysis has examined how pricing, competition, and regulation jointly shape welfare and market outcomes (Bi et al., 2024; Bimpikis et al., 2024; Bi et al., 2025). However, most existing models rely on restrictive assumptions, notably single-round interactions and complete information (CI). Extending these models to scenarios with evolving and incomplete information (INC) is analytically intractable due to the interdependence of decisions across temporal evolution and heterogeneous information structures.

Agent-based modeling (ABM) provides an alternative approach for studying complex environments. Recent work has shown that LLMs can serve as effective agents in ABMs, enabling the simulation of adaptive behaviors in complex social and economic systems (Zhang et al., 2025). Their adaptability further allows the exploration of emergent phenomena and regulatory impacts (Li et al., 2024a; Mou et al., 2025). However, equilibrium-consistent decision-making requires agents to reason over payoffs, beliefs, and strategic dependencies, which are capabilities that LLMs do not inherently possess (Duan et al., 2024). It remains unclear whether LLM-driven agents can reliably approximate equilibrium-consistent decision-making, especially in evolving and INC data marketplaces.

To bridge this gap, we propose **EvoDM**, an

*Corresponding Author (`csqli@comp.polyu.edu.hk`).

Findings of the Association for Computational Linguistics: ACL 2026, pages 32509–32531

ABM framework for evolving and INC data marketplaces. EvoDM extends the static three-layer data marketplace to support temporally evolving interactions and INC. To maintain analytical grounding, we derive tractable equilibrium solutions under simplified regimes of CI, Evolving, and β-INC, serving as analytical benchmarks for evaluating agents' equilibrium-consistent decisions under increasingly realistic conditions. Building upon the EvoDM, we introduce **Datamart-Agent**, an LLM-driven agent with self-evolving strategic capability designed to operate within EvoDM. Datamart-Agent employs role-conditioned prompts to encode objectives, constraints, and reasoning patterns, enabling utility-grounded decisions. A dynamic game tree records the evolving sequence of interactions and expands as new strategies and entrants emerge, while a memory retrieval mechanism identifies semantically similar historical trajectories. These components enable agents to accumulate experience, refine their decisions without parameter updates, and generalize across scenarios.

Our study follows a two-stage objective. The primary objective is to benchmark whether LLM agents can execute equilibrium-consistent decisions in analytically tractable data marketplaces with verifiable ground truth. We verify Datamart-Agent under tractable CI, Evolving, and β-INC settings. It achieves the lowest Utility Gap and is over 20% more Pass@ϵ than strong ABM baselines. The secondary objective is to use the validated agents in EvoDM to study competition and regulation in more realistic settings where exact equilibria are unavailable. We demonstrate that strict data-deletion rules enhance user utility, while dynamic pricing benefits both buyers and platforms. Cost-aware privacy mandates significantly improve fairness, reducing lifespan inequality by over 40% compared to uniform regulations.

Our contributions are summarized as follows:

1. We propose **EvoDM**, a controlled ABM framework for evaluating agent decisions in evolving INC data marketplaces, with equilibrium benchmarks for validation.

2. We introduce **Datamart-Agent**, an LLM-driven agent with self-evolving strategic capability that achieves equilibrium-consistent decision-making in EvoDM.

3. We demonstrate that EvoDM, with the validated Datamart-Agent, supports exploratory analysis of competition and regulation.

2 Related Work

Data marketplaces study pricing, exchange, and regulation among users, platforms, and buyers. Prior research investigates data trading mechanisms (Yu and Zhang, 2017; Koutris et al., 2015; Jia et al., 2019; Liu et al., 2021). These efforts converge on user-platform-buyer interactions (Fallah et al., 2024), with equilibrium analyzes under competition, privacy, and bargaining (Bi et al., 2024, 2025). Recent approaches address valuation and strategic uncertainty (Castro Fernandez, 2022; Amiri et al., 2023; Wang and Jia, 2023), but largely rely on static or CI assumptions, leaving evolving and INC settings underexplored.

Recent work integrates LLMs into ABMs to support language-based reasoning and adaptive behavior beyond rule-based agents (Gao et al., 2024). LLM agents have demonstrated promising performance in social and economic simulations (Huang et al., 2022; Xie et al., 2024; Wu et al., 2024; Li et al., 2024a), yet most focus on behavioral realism rather than rational decisions. This gap motivates our work on equilibrium-consistent decision-making LLM agents for ABM of data marketplaces. A detailed review is provided in Appendix B.

3 Three-Layer Data Marketplace

In this section, we first formalize the static CI three-layer data marketplace, and then theoretically extend it to i) a temporally evolving setting and ii) a buyer-valuation INC setting (β-INC), yielding ground-truth equilibria and utilities to evaluate the reliability of agents in data marketplace ABMs.

3.1 Preliminary: Static CI Marketplace

Unlike classical markets, data marketplaces treat data as an information good whose value depends on signal quality rather than exclusivity (Babaioff et al., 2012), quantify privacy loss as information leakage controlled via noise (Ghosh and Roth, 2011), and exhibit a multi-layer structure with incentives among users, platforms, and buyers (Pei, 2020). Platform-provided information is often substitutable and aggregatable for the buyer under Gaussian models (Koutris et al., 2015). We begin by formalizing the static CI three-layer data marketplace of Fallah et al. (2024).

Definition 1 (Static CI Marketplace). *A static CI three-layer data marketplace is a four-stage game defined as follows:* i) *Each platform $i \in [K]$ simultaneously chooses an **entry** decision $e_i \in \{0,1\}$ and a buyer-facing **noise** level $\sigma_i \geq 0$.* ii) *The user selects a **sharing** decision $\boldsymbol{a} \in \{0,1\}^K$.* iii) *Entered platforms post **prices** $\boldsymbol{p} \in \mathbb{R}_+^K$.* iv) *The data buyer chooses **purchases** $\boldsymbol{b} \in \{0,1\}^K$.*

We next specify the game's payoff structure. Utilities are defined in terms of mutual information, capturing the informational benefits obtained by platforms and the buyer, as well as the privacy leakage incurred by the user.

Definition 2 (Utility). *Let $\boldsymbol{\theta} \sim \mathcal{N}(0, I_d)$ denote the user's data, with platform signals $s_i = \boldsymbol{x}_i^\top \boldsymbol{\theta} + \mathcal{N}(0,1)$ and perturbed signals $\tilde{s}_i = s_i + \mathcal{N}(0, \sigma_i^2)$. The utilities of platform i, the user u, and the buyer b are given by,*

$$\mathcal{U}_i(\boldsymbol{\sigma},\boldsymbol{e},\boldsymbol{a},\boldsymbol{p},\boldsymbol{b}) = [a_i(\mathcal{I}_i + b_i p_i) - c_i] \cdot e_i, \quad (1)$$

$$\mathcal{U}_u(\boldsymbol{\sigma},\boldsymbol{e},\boldsymbol{a},\boldsymbol{p},\boldsymbol{b}) = \sum_i a_i e_i \mathcal{I}_i - \alpha \mathcal{I}(\boldsymbol{\sigma},\boldsymbol{e},\boldsymbol{a},\boldsymbol{b}), \quad (2)$$

$$\mathcal{U}_b(\boldsymbol{\sigma},\boldsymbol{e},\boldsymbol{a},\boldsymbol{p},\boldsymbol{b}) = \beta\, \mathcal{I}(\boldsymbol{\sigma},\boldsymbol{e},\boldsymbol{a},\boldsymbol{b}) - \sum_i b_i p_i, \quad (3)$$

where $\mathcal{I}_i$ denote user side revealed information as formally defined in Prop. A1.

A detailed rationale for the settings of the information framework (*i.e.*, the Gaussian data signal) and utilities is provided in Appendix C.1.

3.2 Equilibrium in the Evolving Marketplace

As real-world data marketplaces involve evolving rounds of interaction among participants, including new entrants joining and incumbents leaving, we extend the static marketplace to T discrete rounds. At each round $t \in [T]$, a new buyer arrives with $(\beta_t, \boldsymbol{y}_t)$, while the active platform set evolves from incumbents S_{t-1} to S_t by admitting a subset of new entrants $S_t^{\text{New}} \subseteq [K_t] \setminus S_{t-1}$. Therefore, the per-round utilities are given by,

Proposition 1 (Per-Round Utilities). *Let $P_{t-1} := \sum_{i \in S_{t-1}} p_i$ denote the total incumbent price carried into round t (prices are set at entry and then fixed). Then for $t = 1, \ldots, T$:*

$$\mathcal{U}_u^{(t)} = 0; \quad (4)$$

$$\mathcal{U}_i^{(t)} = \tfrac{1}{2} + p_i - c_i, \qquad \forall i \in S_{t-1}; \quad (5)$$

$$\mathcal{U}_j^{(t)} = \tfrac{1}{2} + \frac{\beta_t}{2\alpha} - c_j, \qquad \forall j \in S_t^{\text{New}}; \quad (6)$$

$$\mathcal{U}_b^{(t)} = \beta_t \mathcal{I}(\boldsymbol{\sigma}_t) - \Big(P_{t-1} + \frac{\beta_t m_t}{2\alpha}\Big). \quad (7)$$

where $m_t := |S_t^{\text{New}}|$ and $\boldsymbol{\sigma}_t$ is the round-t noise profile. For platforms entering in round $\tau_i \geq 2$, $p_i = \beta_{\tau_i}/(2\alpha)$; for $\tau_i = 1$, p_i follows the static CI equilibrium pricing.

The result demonstrates that cumulative leakage tightens the user's privacy constraint over time, where later entrants must inject larger noise to be accepted, and entry ceases once the implied equilibrium noise becomes infeasible. The detailed proofs could be found in Appendix C.2.

3.3 β-INC: Incomplete Information on Buyer Valuation

The CI data marketplace assumes an ideal situation in which the user and the platforms can observe β. We relax this to a more realistic buyer-valuation INC setting,

Assumption 1 (Private Buyer Valuation). *The buyer's valuation β is drawn from a continuous, atomless, common-knowledge CDF F on $[\underline{\beta}, \overline{\beta}]$ (uniform unless stated otherwise).*

The platforms and the user must form beliefs consistent with a Perfect Bayesian Equilibrium. Let the platform's *per-unit price* be $z_i := \frac{p_i}{\mathcal{I}(\sigma_i)}$. A buyer with value β purchases from platform i iff $\beta \geq z_i$, so the purchase probability is $1 - F(z_i)$. This induces two regimes (*i.e. low-privacy regime* $\alpha \leq \alpha_{\text{crit}}$ and *high-privacy regime* $\alpha > \alpha_{\text{crit}}$) separated by a user privacy threshold α_{crit}, as formally discussed in Appendix C.3. Under β-INC, prices respond to uncertain buyer demand, privacy is enforced on average rather than strictly, and utilities are expectation-based, leading to screening losses and selective market participation.

These equilibria, across various data marketplace situations, provide ground truth for verifying the effectiveness of ABMs in data marketplaces.

4 EvoDM

Real-world data marketplaces evolve under conditions of INC, with platforms, users, and buyers adjusting their decisions in response to changing conditions. While existing three-layer marketplace models capture core strategic interactions, they are largely static and assume full observability, limiting their applicability. To model adaptive behavior under uncertainty, we propose **EvoDM**, a data marketplace ABM framework that incorporates temporal evolution and asymmetric information, as illustrated in Figure 1(a). Platforms i, the user u, and buyers b are modeled as LLM-driven agents in a multi agent Markov game $\mathcal{M} = \langle \mathcal{S}_*, \mathcal{O}_*, \mathcal{A}_*, \mathcal{U}_*, \mathcal{P} \rangle$, where $\mathcal{S}_*$ denotes internal states, $\mathcal{O}_*$ observations, $\mathcal{A}_*$ action spaces, $\mathcal{U}_*$ utilities, and $\mathcal{P}$ marketplace dynamics.

The marketplace unfolds over discrete rounds $t = 1, \ldots, T$, each representing a complete trading cycle. At round t, agents update internal states $\mathcal{S}_*^{(t)}$ and observations $\mathcal{O}_*^{(t)}$, which jointly determine feasible actions $\mathcal{A}_*^{(t)}$. Interactions proceed as follows.

1) Platform Retention and Entry. Incumbent and candidate platforms choose participation via $\mathcal{A}_e^{(t)} = \{e_i^{(t)} \in \{0,1\} \mid i \in S_{t-1} \cup S_t^{\text{new}}\}$. Each incumbent $i \in S_{t-1}$ evaluates its private state $\mathcal{S}_i^{(t)} = \{\boldsymbol{x}_i, c_i, b_i^{(t)}\}$ and either remains active with $e_i^{(t)} = 1$ or exits permanently with $e_i^{(t)} = 0$ to avoid future operational costs c_i. Each candidate $j \in S_t^{\text{new}}$ samples $\mathcal{S}_j^{(t)} = \{\boldsymbol{x}_j, c_j\}$ with $|\boldsymbol{x}_j| = 1$, observes public information $\mathcal{O}_j^{(t)} = \{\sigma_i, p_i^{(t-1)} \mid i \in S_{t-1}\}$, and decides whether to enter.

2) Platform Offer. Each newly entered platform selects a privacy noise level via $\mathcal{A}_\sigma^{(t)} = \{\sigma_j \in \mathbb{R}^+ \mid j \in S_t^{\text{new}}\}$ based on its private state $\mathcal{S}_j^{(t)} = \{\boldsymbol{x}_j, c_j\}$ and public observation $\mathcal{O}_j^{(t)}$. The chosen σ_j remains fixed throughout the platform lifetime, ensuring consistent privacy guarantees.

3) User Data Sharing. Given active platforms and announced privacy levels, the user with internal state $\mathcal{S}_u^{(t)} = \{\theta, \alpha\}$ observes $\mathcal{O}_u^{(t)} = \{\sigma_i^{(t)} \mid i \in S_t\}$ and selects $\mathcal{A}_u^{(t)} = \{a_i^{(t)} \in \{0,1\} \mid i \in S_t,\ e_i^{(t)} = 1\}$. Under the *strict deletion* regime with deletion-on-withdrawal (Commission, 2012), the current choice applies to all active platforms, and any previously selected platform not chosen in round t must delete stored data. Under the *legacy retention* regime, previously selected platforms retain data while active, and $\boldsymbol{a}^{(t)}$ primarily concerns new entrants. The two regimes differ only in feasible actions and state transition rules.

4) Platform Pricing. Platforms that both enter and receive data set prices using internal state $\mathcal{S}_i^{(t)} = \{\boldsymbol{x}_i, \boldsymbol{s}_i, \sigma_i\}$ and observation $\mathcal{O}_i^{(t)} = \{p_i^{(t-1)}, e^{(t-1)}\}$ via $\mathcal{A}_p^{(t)} = \{p_i^{(t)} \in \mathbb{R}^+ \mid i \in S_t,\ e_i^{(t)} = a_i^{(t)} = 1\}$. We consider two regimes: a *price persistence* regime, in which prices evolve gradually, and a *price reoptimization* regime, in which prices are allowed to adjust at each round. The platforms cannot observe the private state of other agents as in the real-world.

5) Buyer Selection. A buyer with internal state $\mathcal{S}_b^{(t)} = \{\boldsymbol{y}_t, \beta_t\}$ observes public offers $\mathcal{O}_b^{(t)} = \{\sigma_i, p_i^{(t)} \mid i \in S_t,\ e_i^{(t)} = a_i^{(t)} = 1\}$ and selects $\mathcal{A}_b^{(t)} = \{b_i^{(t)} \in \{0,1\} \mid i \in S_t,\ e_i^{(t)} = a_i^{(t)} = 1\}$. For each purchase, the buyer receives $\tilde{s}_i^{(t)} = \boldsymbol{s}_i^{(t)} + \eta_i^{(t)}$, where $\eta_i^{(t)} \sim \mathcal{N}(0, \sigma_i^2)$, without observing true signals or platform costs.

All agents receive per round utilities $\mathcal{U}_*^{(t)}$ as defined in Section 3, with cumulative utilities $\mathcal{U}_*^{(1:t)} = \sum_{t'=1}^{t} \mathcal{U}_*^{(t')}$. Platforms with $\mathcal{U}_i^{(1:t)} < 0$ exit endogenously. The EvoDM enables systematic analysis of data marketplaces under evolving and fully INC settings. Appendix C.4 provides theoretical justification for the modeling of EvoDM.

5 Datamart-Agent

As shown in Figure 1(b), **Datamart-Agent** is an LLM-driven agent designed to replicate equilibrium-consistent decisions in simulating data marketplaces within EvoDM. It integrates three components that jointly enable agents to perceive the marketplace, accumulate experience, and iteratively refine decisions toward equilibrium-consistent behaviors.

5.1 Dynamic Role-Conditioned Perception

To model heterogeneous cognition in data marketplaces, Datamart-Agent employs a dynamic role-conditioned perception that translates structured numerical states and observations into natural-language descriptions tailored to each role. Participants receive role-specific perceptions that combine private internal states with publicly observable market signals, ensuring consistency with the marketplace's information constraints.

Perceptions are encoded into compact natural-language representations that translate numerical states and symbolic market dynamics into semantically meaningful descriptions. These prompts capture pricing, privacy levels, competition, and cooperation, and are updated each round to reflect changes such as platform entry, exit, and user sharing decisions. By bridging structured perception with linguistic reasoning, this module enables agents to interpret uncertainty and evolving incentives without relying on fixed decision rules.

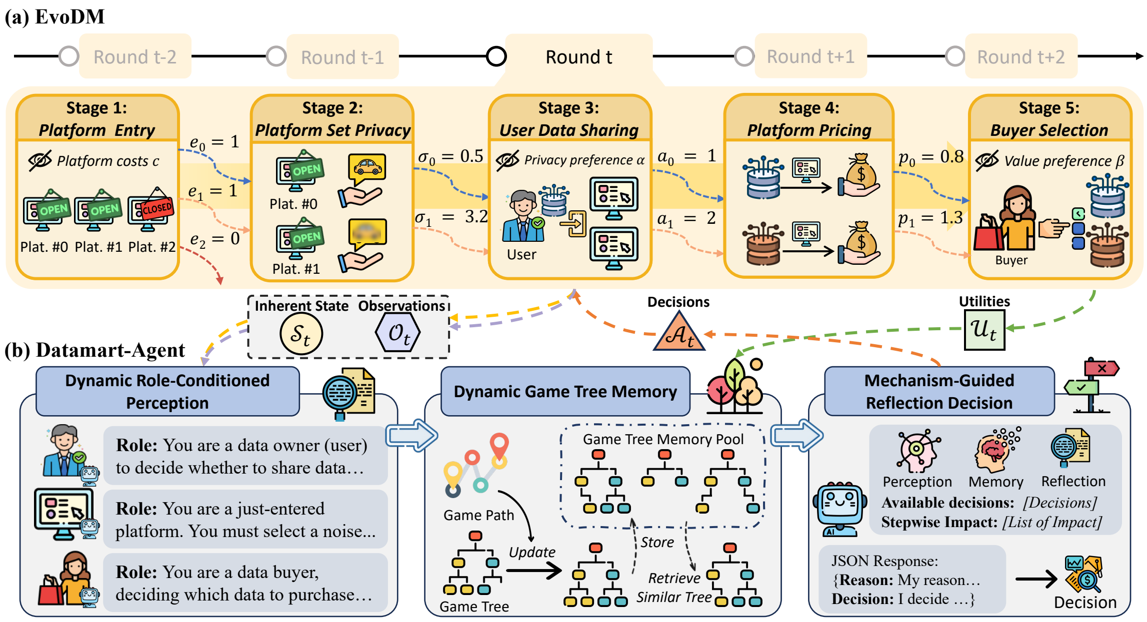

Figure 1: The framework of EvoDM and Datamart-Agent.

The perception prompt structure is provided in Appendix D.

5.2 Dynamic Game Tree Memory

Decision-making in data marketplaces is sequential and interdependent. To support long-horizon reasoning, Datamart-Agent maintains a dynamic game tree memory that incrementally organizes historical interactions into structured decision trees (Hua et al., 2024). A sequential decision game is represented as a game tree

$$T = (\mathcal{V}, \mathcal{E}),\ \mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}, \tag{8}$$

where nodes correspond to partial action sequences and leaf nodes store realized utilities.

Unlike static game trees, the memory grows online as the marketplace evolves. For each role $* \in \{u, i, b\}$, the agent maintains a memory pool $\mathcal{T}_*$ indexed by role-specific state-observation pairs. When a new round begins, the agent initializes a new tree if no matching index exists, or expands an existing tree by appending the current action and its observed utility. This process consolidates experience across rounds, enabling generalization beyond one-shot reasoning.

When making new decisions, the module performs similarity-based *retrieval* to recall relevant past experiences. Given the current state-observation pair $\#_*^{(t)}$ and each stored index $\#_{*,j}$, the similarity is computed as,

$$\mathrm{sim}\left(\#_*^{(t)}, \#_{*,j}\right) = \frac{\langle \#_*^{(t)}, \#_{*,j} \rangle}{\|\#_*^{(t)}\| \, \|\#_{*,j}\|}. \tag{9}$$

The top-N most similar trees, $\{T_{*,j}\}_{j=1}^{N}$, are retrieved and summarized into a compact natural-language reflection that highlights historical decisions and their corresponding payoffs. This reflection serves as contextual evidence for the agent, helping it anticipate the outcomes of alternative strategies and maintain equilibrium consistency across rounds. The game tree memory prompt is detailed in Appendix D.

5.3 Mechanism-Guided Reflection Decision

The mechanism-guided reflection decision enables each agent to reason about the consequences of its actions through explicit utility formulations, allowing for more equilibrium-consistent decision-making. Rather than relying only on pattern matching, the agent is prompted to evaluate how candidate actions affect its utility under the market mechanism in Section 3. The prompt specifies the payoff structure and decision constraints for the current role, but it does not directly provide the optimal action, as illustrated in Appendix D. In evolving, incomplete-information settings, effective decisions still require the agent to infer the hidden states of other participants and anticipate how future actions may change the realized utility. By combining the mechanism-guided reflection with prior perception and memory, the agent's actions align with equilibrium-consistent behavior rather than heuristic imitation, ensuring that decisions remain both context-aware and utility-driven within the EvoDM.

At each round t, the agent integrates its perception, memory, and mechanism awareness into a unified decision context. Specifically, we construct,

$$\mathcal{C}_*^{(t)} = \Big(\phi_{\text{perc}}(\mathcal{S}_*^{(t)},\mathcal{O}_*^{(t)}),\phi_{\text{mem}}(\mathcal{T}_*,\#_*^{(t)}),\phi_{\text{ref}}(\mathcal{U}_*,\mathcal{M})\Big), \quad (10)$$

where ϕ_{perc} denotes the role-conditioned perception that maps internal states and public observations to a natural-language representation, ϕ_{mem} retrieves relevant experience from the dynamic game tree memory $\mathcal{T}_*$ based on the current role-specific identifier $\#_*^{(t)}$, and ϕ_{ref} embeds the utility formulation and equilibrium constraints induced by the marketplace mechanism $\mathcal{M}$. The final decision is,

$$a_*^{(t)} = \text{LLM}\Big(\mathcal{C}_*^{(t)}\Big), \quad \text{s.t. } a_*^{(t)} \in \mathcal{A}_*^{(t)}. \quad (11)$$

This formulation allows Datamart-Agent to combine mechanism information with role-specific context and retrieved experience, producing structured and utility-aware decisions under evolving and INC market conditions.

6 Experiments

In this section, we conduct experiments to study the simulation performance of the Datamart-Agent in the EvoDM. Our evaluation is organized into two stages. Section 6.2 focuses on quantitative validation against analytical ground truth in tractable settings. Sections 6.3 and 6.4 then use the validated agents for simulation-based analysis in settings where exact equilibrium benchmarking is unavailable or intentionally relaxed. The experiments are structured around three core questions:

RQ1 (Effectiveness): To what extent can the proposed Datamart-Agent accurately replicate equilibrium behaviors within the EvoDM?

RQ2 (Emergent Dynamics): How do agents adapt and interact under evolving and INC data marketplaces, and what emergent phenomena arise in such realistic multi-platform environments?

RQ3 (Application): Can the Datamart-Agent support practical applications?

6.1 Experimental Setup

We evaluate the Datamart-Agent on benchmarks derived from our EvoDM under three settings: **CI**, **Evolving**, and **β-INC**, each with analytically tractable equilibria. We randomly generate data marketplace scenarios for training, validation, and testing, and additionally construct out-of-distribution (OOD) scenarios with shifted parameter ranges and numbers of competing platforms to evaluate generalization. Detailed benchmark construction is provided in Appendix A.1.

In our experiments, we compare Datamart-Agent with both RL-based and LLM-based baselines on EvoDM. RL agents learn policies through interaction with the environment, whereas LLM agents make decisions by conditioning on role descriptions, observations, and mechanism information provided through prompting. Our baselines include: **RL-ABM** (Lowe et al., 2017), and LLM agents: **LLM-State** (Zhang et al., 2024b), **LLM-Recent** (Zhang et al., 2024c; Piao et al., 2025), **LLM-I** (Li et al., 2024a), and **LLM-E** (Hua et al., 2024). Details are illustrated in Appendix A.2.

We evaluate ABM performance in the EvoDM using i) Utility Gap $\Delta\mathcal{U}$ and ii) Pass@ε, as described in Appendix A.3. The backbone models and hyperparameter settings are illustrated in Appendix A.4.

6.2 Effectiveness in Simulation (RQ1)

6.2.1 Main Results

To verify the effectiveness of our Datamart-Agent in making equilibrium-consistent decisions and achieving equilibrium utilities in the data marketplace ABM framework, we evaluate it against RL-based and LLM-driven baselines across six data marketplace settings, as summarized in Table 1. The results show that Datamart-Agent most accurately replicates equilibrium behaviors in EvoDM. It consistently achieves the lowest equilibrium utility gap and the highest Pass@ϵ across all evaluated settings, indicating strong alignment with equilibrium-consistent decisions under evolving and INC settings. When instantiated with GPT-4o, Datamart-Agent exceeds the best LLM-E by over 20% in Pass@ϵ, demonstrating its effectiveness in making equilibrium-consistent decisions.

Figure 2 further confirms this effectiveness at the decision level. Datamart-Agent attains the highest accuracy for discrete equilibrium decisions and the lowest MSE for continuous decisions such as privacy noise σ and price p, outperforming all baselines. These results indicate that Datamart-Agent not only matches equilibrium outcomes in aggre-

Agents	Models	CI		CI-OOD		Evolving		Evolving-OOD		β-INC		β-INC-OOD	
		$\Delta\mathcal{U}(\downarrow)$	Pass@$\epsilon(\uparrow)$	$\Delta\mathcal{U}(\downarrow)$	Pass@$\epsilon(\uparrow)$	$\Delta\mathcal{U}(\downarrow)$	Pass@$\epsilon(\uparrow)$	$\Delta\mathcal{U}(\downarrow)$	Pass@$\epsilon(\uparrow)$	$\Delta\mathcal{U}(\downarrow)$	Pass@$\epsilon(\uparrow)$	$\Delta\mathcal{U}(\downarrow)$	Pass@$\epsilon(\uparrow)$
RL-MAB	DQN	0.354	0.200	0.370	0.182	0.385	0.158	0.410	0.123	0.270	0.417	0.466	0.223
	PPO	0.344	0.251	0.360	0.236	0.375	0.202	0.400	0.185	**0.125**	0.530	0.292	0.307
LLM-State	LLaMA3.1-8B	0.302	0.395	0.347	0.306	0.301	0.313	0.429	0.208	0.334	0.395	0.401	0.297
	Qwen2.5-32B	0.285	0.415	0.332	0.325	0.285	0.335	0.395	0.225	0.301	0.411	0.348	0.345
	GPT-4o	0.225	0.570	0.315	0.331	0.235	0.450	0.335	0.350	0.256	0.485	0.302	0.415
LLM-Recent	LLaMA3.1-8B	0.351	0.336	0.359	0.276	0.359	0.259	0.428	0.215	0.405	0.314	0.467	0.262
	Qwen2.5-32B	0.304	0.356	0.331	0.303	0.333	0.264	0.413	0.231	0.353	0.359	0.398	0.327
	GPT-4o	0.264	0.513	0.311	0.343	0.282	0.337	0.351	0.349	0.335	0.412	0.346	0.378
LLM-I	LLaMA3.1-8B	0.173	0.632	0.204	0.525	0.177	0.599	0.382	0.414	0.323	0.411	0.402	0.318
	Qwen2.5-32B	0.169	0.662	0.196	0.582	0.165	0.630	0.326	0.434	0.288	0.467	0.285	0.410
	GPT-4o	0.147	0.701	0.179	0.589	0.160	0.675	0.313	0.455	0.256	0.491	0.240	0.524
LLM-E	LLaMA3.1-8B	0.182	0.615	0.185	0.666	0.189	0.582	0.175	0.565	0.268	0.525	0.357	0.453
	Qwen2.5-32B	0.177	0.659	0.183	0.649	0.165	0.652	0.165	0.594	0.242	0.552	0.274	0.481
	GPT-4o	<u>0.146</u>	<u>0.742</u>	<u>0.144</u>	<u>0.787</u>	<u>0.154</u>	<u>0.702</u>	<u>0.158</u>	<u>0.632</u>	0.213	<u>0.661</u>	<u>0.236</u>	<u>0.602</u>
Ours	LLaMA3.1-8B	0.193	0.603	0.188	0.651	0.165	0.674	0.178	0.588	0.277	0.440	0.288	0.470
	Qwen2.5-32B	0.178	0.652	0.174	0.674	0.161	0.681	0.168	0.597	0.231	0.602	0.264	0.473
	GPT-4o	**0.132**	**0.849**	**0.136**	**0.796**	**0.145**	**0.734**	**0.151**	**0.656**	<u>0.203</u>	**0.712**	**0.215**	**0.681**

Table 1: Main results comparing different RL-based and LLM-driven agents across six data marketplace settings.

gate utility but also reliably reproduces the underlying equilibrium decision rules.

6.2.2 Ablation Study

To evaluate the effectiveness of each module, we conduct ablation studies on the EvoDM benchmarks using GPT-4o for Datamart-Agent. As shown in Table 2, removing any major module consistently degrades equilibrium replication performance. The most significant drop occurs when removing the natural language representation in the *Dynamic Role-Conditioned Perception*, which increases the equilibrium utility gap by 0.391, indicating that semantic grounding is critical for accurate equilibrium alignment. For the *Dynamic Game Tree Memory*, removing similarity-based retrieval leads to substantial degradation, whereas removing tree memory alone results in only a minor change, showing that experience retrieval is essential for maintaining equilibrium consistency. In addition, removing the *Mechanism Guided Reflection Decision* increases the utility gap, confirming that explicit mechanism guidance enhances equilibrium-consistent decision-making.

We also examine the computational overhead introduced by the memory module. Although the dynamic game tree expands online, inference only uses the Top-N retrieved paths rather than the full tree, which keeps the prompt length bounded. Table 3 reports the average token usage and inference latency over each scenario in the CI setting with Qwen-2.5-32B-Instruction on one Nvidia H20 GPU. The additional overhead remains modest compared with strong LLM-agent baselines, suggesting that our game tree memory module is computationally manageable.

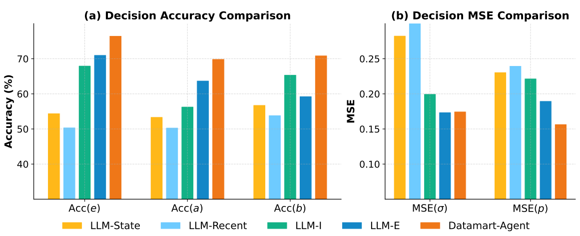

Figure 2: Agent Decision Performance.

Modules	CI	Evolving	β-INC	Avg. Chg.
Datamart-Agent	0.132	0.145	0.203	—
Perception				
w/o *Natural Language Representation*	0.525	0.532	0.595	+0.391
Memory				
w/o *Tree Memory*	0.136	0.151	0.225	+0.011
w/o *Similar Retrieval*	0.278	0.290	0.341	+0.143
w/o *Tree Memory + Similar Retrieval*	0.282	0.296	0.340	+0.146
Action				
w/o *Mechanism-Guided Reflection*	0.139	0.163	0.254	+0.025

Table 2: Ablation results of Datamart-Agent on $\Delta\mathcal{U}$ ($\downarrow$).

	LLM-State	LLM-Recent	LLM-I	LLM-E	Ours
Avg. Tokens	7232	8322	9318	9136	9778
Avg. Time (s)	13.5	15.7	17.6	17.2	18.8

Table 3: Computational overhead in CI scenarios.

We further analyze similarity-based retrieval in the most challenging β-INC-OOD setting. We log the Top-N retrieval similarity at each decision step and summarize the resulting $\Delta\mathcal{U}$ by similarity bin. Table 4 shows that higher retrieval similarity is associated with smaller $\Delta\mathcal{U}$, with a Pearson correlation of -0.331. This result further supports the importance of retrieving trajectories close to the current data marketplace context.

Similarity	[0.00, 0.25)	[0.25, 0.50)	[0.50,0.75)	[0.75, 1.00]
$\Delta\mathcal{U}$ (↓)	0.362	0.295	0.271	0.254

Table 4: $\Delta\mathcal{U}$ *vs.* retrieval similarity over β-INC-OOD.

6.3 Emergent Market Dynamics (RQ2)

Building on the verified effectiveness of Datamart-Agent in aligning equilibria, we examine the emergent behaviors of the Datamart-Agent in the EvoDM to provide simulation-driven insights. All simulations employ the GPT-4o as the backbone.

6.3.1 Utilities under Different Mechanisms

We examine how marketplace mechanisms influence utility distributions across users, platforms, and buyers by varying two core design dimensions: i) *strict deletion* versus *legacy retention*, and ii) *price persistence* versus *price re-optimization*. Simulations are repeated 50 times with $T = 25$ and $K_{\max} = 8$, and aggregate utilities are reported in Figure 3. Strict deletion consistently improves user utility by allowing adaptive withdrawal and limiting cumulative information leakage, while slightly reducing platform and buyer utilities due to constrained long-term data access. In contrast, price re-optimization increases buyer utility and moderately benefits platforms through competitive pricing, but slightly reduces user utility as higher trading intensity amplifies the loss of privacy. These results illustrate clear trade-offs induced by regulatory and pricing mechanisms, demonstrating how agent behaviors and utilities co-evolve under INC.

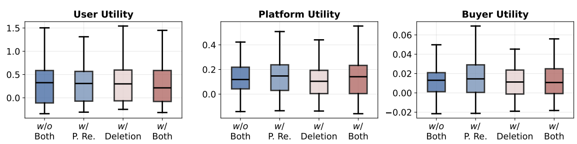

Figure 3: Utilities under different mechanisms.

6.3.2 Behavioral Responses to Platform

We study how platform privacy and pricing jointly affect market participation on both the user and buyer sides. User sharing behavior is analyzed as a function of platform privacy levels under heterogeneous privacy preferences. Specifically, we group users into low ($\alpha \in [0, 1)$), mid ($\alpha \in [1, 4)$), and high ($\alpha \in [4, 5]$) according to their privacy sensitivity. As shown in Figure 4 (a), privacy sensitivity has a strong impact on user participation. User sharing probability increases with stronger privacy

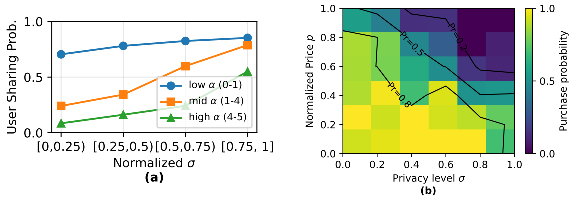

Figure 4: Behavioral responses to platform privacy and pricing. (a) User sharing probability as a function of privacy level under heterogeneous privacy preferences. (b) Buyer purchase probability across price and privacy combinations, illustrating the price-privacy trade-off.

protection across all groups, with low-α users sharing consistently while high-α users respond sharply only at higher privacy levels.

We uniformly sample platforms and buyers to measure the purchase probability of buyers across different combinations of normalized price p and privacy level σ. As shown in Figure 4 (b), the buyer's purchase probability decreases with higher prices and increases with stronger privacy protection, revealing a clear price-privacy trade-off. Notably, high privacy levels can partially offset higher prices, indicating that privacy acts as a complementary value signal in buyer decision-making.

6.3.3 Lifespan over Heterogeneous Platforms

We analyze platform survival under varying market horizons and platform density by measuring normalized platform lifespan. With $K_{\max} = 16$, increasing the horizon from $T = 10$ to $T = 50$ raises the normalized lifespan of low-cost platforms ($c_i \leq 0.3$) from approximately 0.15 to 0.25, while high-cost platforms ($c_i \geq 0.8$) exhibit near-zero lifespan across all horizons. Fixing $T = 25$ and increasing $K_{\max}$ from 8 to 32 reduces the normalized lifespan of low-cost platforms from around 0.30 to below 0.15, indicating intensified competition. These patterns, shown in Figure 5, reveal emergent selection effects where competitive pressure drives platform exit and market turnover.

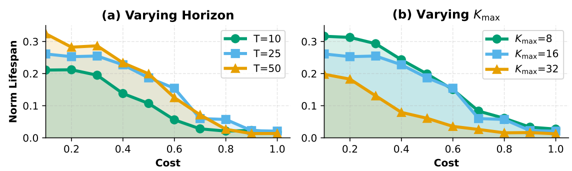

Figure 5: Platform lifespan over varying competitions.

6.3.4 Preference Heterogeneity and Market Composition

We discuss how heterogeneous user privacy preferences interact with market composition to shape platform selection behavior. Concretely, platforms are categorized as privacy-first with high privacy noise $\sigma = 0.9$ or utility-first with low privacy noise $\sigma = 0.1$. The market composition is controlled by the *high-σ platform ratio*, defined as the fraction of active platforms that are privacy-first. Users are grouped into low-, mid-, and high-α cohorts and evaluated under the legacy retention regime.

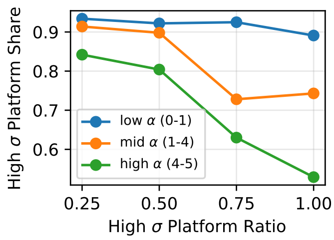

Figure 6: Share of high-σ platforms selected by users with different privacy preferences under varying high-σ platform ratios.

As shown in Figure 6, low-α users consistently maintain a high probability of sharing with high-σ platforms across all market compositions, indicating weak sensitivity to changes in platform availability. In contrast, high-α users exhibit a marked decline in the share of high-σ platforms they engage with as the high-σ platform ratio increases. This pattern reflects increased selectivity among privacy-sensitive users, who concentrate their sharing on a limited subset of trusted platforms when privacy-first options become abundant, rather than distributing interactions uniformly. The results demonstrate how heterogeneous privacy preferences lead to differentiated platform selection and concentration effects, highlighting nontrivial interactions between user behavior and market composition in data marketplaces.

6.4 Case Study: Privacy Regularities (RQ3)

Motivated by the market patterns observed in Section 6.3, we next study how privacy regulation affects platform lifespan fairness under cost heterogeneity. We compare representative regulatory regimes to examine which policy designs better balance fairness and platform viability in dynamic data marketplaces. We quantify fairness using the Gini coefficient (Gini, 1921),

$$G_L = \frac{\sum_{i=1}^{K} \sum_{j=1}^{K} |L_i - L_j|}{2K \sum_{i=1}^{K} L_i}, \qquad (12)$$

where L_i denotes the lifespan of platform i, and lower G_L indicates more balanced lifespans.

Regime	$c_i \in (0.0, 0.4)$	$c_i \in [0.4, 0.7)$	$c_i \in [0.7, 1.0]$	**Gini**($\downarrow$)
None	0.248	0.164	0.052	0.256
Ban	0.121	0.033	0.011	0.435
Uniform	0.198	0.086	0.028	0.387
Tiered	0.117	0.042	0.039	0.259

Table 5: Lifespan distribution and Gini coefficients under different regulatory regimes.

We compare three representative regulatory regimes: i) a "Privacy Ban" with $\underline{\sigma}_i = \infty$, ii) a "Uniform Minimum Privacy Mandate" with $\underline{\sigma}_i = 0.5$, and iii) a "Tiered Privacy Mandate" that applies $\underline{\sigma}_i = \infty$ to platforms with $c_i < 0.5$ and $\underline{\sigma}_i = 0.5$ otherwise. Table 5 reports the resulting lifespan distributions and fairness outcomes. The privacy ban sharply reduces platform participation and increases inequality, resulting in a Gini coefficient of 0.435. The uniform minimum privacy mandate also exacerbates imbalance, with a Gini coefficient of 0.387, disproportionately harming high-cost platforms. In contrast, the tiered privacy mandate significantly improves fairness, reducing the Gini coefficient to 0.259 while preserving platform viability across cost groups. It provides insights and feedback for designing privacy protection regulations in data marketplaces.

7 Conclusion

In this paper, we propose EvoDM, an ABM framework for evolving and INC data marketplaces, together with Datamart-Agent, an LLM-driven agent for equilibrium-consistent decision-making. Using analytically tractable CI, Evolving, and β-INC settings, we show that Datamart-Agent more closely matches theoretical equilibria than RL-based and LLM-based baselines. Building on this validation, we further use EvoDM with Datamart-Agent to study competition and privacy regulation in dynamic data marketplaces, revealing platform lifespan inequality under cost heterogeneity and the fairness benefits of cost-aware privacy mandates over uniform regulation.

Limitations

Although the current framework effectively models equilibrium-consistent decision-making and regulatory dynamics, several promising directions remain for future exploration. i) *Persona integration:* incorporating diverse agent personas with varying cognitive styles, risk preferences, and behavioral biases can enrich heterogeneity and enable the simulation of more human-like interactions among participants. ii) *Optimization-based regulation:* extending the current regulator to a multi-objective optimization paradigm can support the automatic discovery of Pareto-efficient policy configurations that balance privacy, fairness, and welfare objectives. Coupling this regulator with feedback signals from the data marketplace ABM and Datamart-Agent would facilitate adaptive policy refinement and closed-loop governance. iii) *Evaluation scope*. The current evaluation focuses on equilibrium-consistent decision execution in analytically tractable settings. It demonstrates reliable decision execution under this benchmark protocol, while it does not consider agent behavior under strategic disequilibrium, against non-equilibrium opponents, or in environments with equilibrium multiplicity.

Acknowledgments

The research described in this paper has been partially supported by the General Research Funds from the Hong Kong Research Grants Council (project No. PolyU 15207322, 15200023, 15206024, and 15224524), Hong Kong Research Grants Council's Theme-based Research Scheme (No. T43-513/23-N), Hong Kong Research Grants Council's Research Impact Fund (No. R1015-23), Hong Kong Research Grants Council's Collaborative Research Fund (No. C1043-24GF), Internal research funds from Hong Kong Polytechnic University (project no. P0059586, P0042693, P0048625, and P0051361), and Sheertek International (HK) Limited. This work was supported by computational resources provided by The Centre for Large AI Models (CLAIM) of The Hong Kong Polytechnic University.

References

2025. Aws data exchange. `https://aws.amazon.com/data-exchange/`.

Melika Abolhassani, Hossein Esfandiari, Mohammad-Taghi Hajiaghayi, Brendan Lucier, and Hadi Yami. 2017. Market pricing for data streams. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31.

Anish Agarwal, Munther Dahleh, and Tuhin Sarkar. 2019. A marketplace for data: An algorithmic solution. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 701–726.

Armen Aghajanyan, Lili Yu, Alexis Conneau, Wei-Ning Hsu, Karen Hambardzumyan, Susan Zhang, Stephen Roller, Naman Goyal, Omer Levy, and Luke Zettlemoyer. 2023. Scaling laws for generative mixed-modal language models. In *International Conference on Machine Learning*, pages 265–279. PMLR.

Mohammad Mohammadi Amiri, Frederic Berdoz, and Ramesh Raskar. 2023. Fundamentals of task-agnostic data valuation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 9226–9234.

Rossella Argenziano and Alessandro Bonatti. 2023. Data markets with privacy-conscious consumers. In *AEA Papers and Proceedings*, volume 113, pages 191–196. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Moshe Babaioff, Robert Kleinberg, and Renato Paes Leme. 2012. Optimal mechanisms for selling information. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 92–109.

Dirk Bergemann, Alessandro Bonatti, and Alex Smolin. 2018. The design and price of information. *American economic review*, 108(1):1–48.

Aditya Bhaskara, Sreenivas Gollapudi, Sungjin Im, Kostas Kollias, Kamesh Munagala, and Govind S Sankar. 2024. Data exchange markets via utility balancing. In *Proceedings of the ACM Web Conference 2024*, pages 57–65.

Yuran Bi, Jinfei Liu, Kui Ren, Yihang Wu, and Yang Cao. 2025. Bargaining-based data markets. In *2025 IEEE 41st International Conference on Data Engineering (ICDE)*, pages 1456–1469. IEEE.

Yuran Bi, Jinfei Liu, Chen Zhao, Junyi Zhao, Kui Ren, and Li Xiong. 2024. Share: Stackelberg-nash based data markets. In *2024 IEEE 40th International Conference on Data Engineering (ICDE)*, pages 3573–3586. IEEE.

Kostas Bimpikis, Ilan Morgenstern, and Daniela Saban. 2024. Data tracking under competition. *Operations Research*, 72(2):514–532.

Christine L Borgman. 2012. The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63(6):1059–1078.

Raul Castro Fernandez. 2022. Protecting data markets from strategic buyers. In *Proceedings of the 2022 International Conference on Management of Data*, pages 1755–1769.

Raul Castro Fernandez. 2023. Data-sharing markets: model, protocol, and algorithms to incentivize the formation of data-sharing consortia. *Proceedings of the ACM on Management of Data*, 1(2):1–25.

Lingjiao Chen, Paraschos Koutris, and Arun Kumar. 2019. Towards model-based pricing for machine learning in a data marketplace. In *Proceedings of the 2019 international conference on management of data*, pages 1535–1552.

Zhengyu Chen, Siqi Wang, Teng Xiao, Yudong Wang, Shiqi Chen, Xunliang Cai, Junxian He, and Jingang Wang. 2025a. Revisiting scaling laws for language models: The role of data quality and training strategies. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 23881–23899.

Zhixun Chen, Zijing Shi, Yaodong Yang, Meng Fang, and Yali Du. 2025b. Hierarchical multi-agent framework for dynamic macroeconomic modeling using large language models. In *Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems*, pages 2460–2462.

European Commission. 2012. Proposal for a regulation of the european parliament and of the council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (general data protection regulation). Procedure 2012/0011(COD).

Thomas M Cover. 1999. *Elements of information theory*. John Wiley & Sons.

Jinhao Duan, Renming Zhang, James Diffenderfer, Bhavya Kailkhura, Lichao Sun, Elias Stengel-Eskin, Mohit Bansal, Tianlong Chen, and Kaidi Xu. 2024. Gtbench: Uncovering the strategic reasoning capabilities of llms via game-theoretic evaluations. *Advances in Neural Information Processing Systems*, 37:28219–28253.

Meta FAIR, Anton Bakhtin, Noam Brown, Emily Dinan, Gabriele Farina, Colin Flaherty, Daniel Fried, Andrew Goff, Jonathan Gray, Hengyuan Hu, and 1 others. 2022. Human-level play in the game of diplomacy by combining language models with strategic reasoning. *Science*, 378(6624):1067–1074.

Alireza Fallah, Michael Jordan, and Ali Makhdoumi. 2024. On three-layer data markets. In *Agentic Markets Workshop at ICML*.

Wenqi Fan, Yao Ma, Qing Li, Yuan He, Eric Zhao, Jiliang Tang, and Dawei Yin. 2019. Graph neural networks for social recommendation. In *The world wide web conference*, pages 417–426.

Raul Castro Fernandez, Pranav Subramaniam, and Michael J Franklin. 2020a. Data market platforms: Trading data assets to solve data problems. *Proceedings of the VLDB Endowment*, 13(11).

Raul Castro Fernandez, Pranav Subramaniam, and Michael J Franklin. 2020b. Data market platforms: trading data assets to solve data problems. *Proceedings of the VLDB Endowment*, 13(12):1933–1947.

Chen Gao, Xiaochong Lan, Nian Li, Yuan Yuan, Jingtao Ding, Zhilun Zhou, Fengli Xu, and Yong Li. 2024. Large language models empowered agent-based modeling and simulation: A survey and perspectives. *Humanities and Social Sciences Communications*, 11(1):1–24.

Arpita Ghosh and Aaron Roth. 2011. Selling privacy at auction. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 199–208.

Corrado Gini. 1921. Measurement of inequality of incomes. *The Economic Journal*, 31(121):124–126.

Denizalp Goktas. 2022. An algorithmic theory of markets and their application to decentralized markets. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 12878–12879.

Qwen Group. 2025. Qwen2. 5-vl technical report. *arXiv preprint arXiv:2412.15115*.

Tiancheng Hu and Nigel Collier. 2024. Quantifying the persona effect in LLM simulations. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 10289–10307, Bangkok, Thailand. Association for Computational Linguistics.

Wenyue Hua, Ollie Liu, Lingyao Li, Alfonso Amayuelas, Julie Chen, Lucas Jiang, Mingyu Jin, Lizhou Fan, Fei Sun, William Wang, and 1 others. 2024. Game-theoretic llm: Agent workflow for negotiation games. *arXiv preprint arXiv:2411.05990*.

Wenlong Huang, Pieter Abbeel, Deepak Pathak, and Igor Mordatch. 2022. Language models as zero-shot planners: Extracting actionable knowledge for embodied agents. In *International conference on machine learning*, pages 9118–9147. PMLR.

Zezhou Huang, Jiaxiang Liu, Daniel Gbenga Alabi, Raul Castro Fernandez, and Eugene Wu. 2023. Saibot: A differentially private data search platform. *Proceedings of the VLDB Endowment*, 16(11):3057–3070.

Nicole Immorlica, Ian A Kash, and Brendan Lucier. 2021. Buying data over time: Approximately optimal strategies for dynamic data-driven decisions. In *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, pages 77–1. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

Ruoxi Jia, David Dao, Boxin Wang, Frances Ann Hubis, Nezihe Merve Gurel, Bo Li, Ce Zhang, Costas Spanos, and Dawn Song. 2019. Efficient task-specific data valuation for nearest neighbor algorithms. *Proceedings of the VLDB Endowment*, 12(11):1610–1623.

Zhuohang Jiang, Pangjing Wu, Ziran Liang, Peter Q Chen, Xu Yuan, Ye Jia, Jiancheng Tu, Chen Li, Peter HF Ng, and Qing Li. 2025a. Hibench: Benchmarking llms capability on hierarchical structure reasoning. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*, pages 5505–5515.

Zhuohang Jiang, Pangjing Wu, Xu Yuan, Wenqi Fan, and Li Qing. 2025b. Qa-dragon: Query-aware dynamic rag system for knowledge-intensive visual question answering. In *2025 KDD Cup Workshop for Multimodal Retrieval Augmented Generation*.

Paraschos Koutris, Prasang Upadhyaya, Magdalena Balazinska, Bill Howe, and Dan Suciu. 2015. Query-based data pricing. *Journal of the ACM (JACM)*, 62(5):1–44.

Nian Li, Chen Gao, Mingyu Li, Yong Li, and Qingmin Liao. 2024a. Econagent: Large language model-empowered agents for simulating macroeconomic activities. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15523–15536.

Yifan Li, Xiaohui Yu, and Nick Koudas. 2021. Data acquisition for improving machine learning models. *Proceedings of the VLDB Endowment*, 14(10):1832–1844.

Zitao Li, Bolin Ding, Liuyi Yao, Yaliang Li, Xiaokui Xiao, and Jingren Zhou. 2024b. Performance-based pricing for federated learning via auction. *Proceedings of the VLDB Endowment*, 17(6):1269–1282.

Jinfei Liu, Jian Lou, Junxu Liu, Li Xiong, Jian Pei, and Jimeng Sun. 2021. Dealer: An end-to-end model marketplace with differential privacy. *Proceedings of the VLDB Endowment*, 14(6).

Ryan Lowe, Yi I Wu, Aviv Tamar, Jean Harb, OpenAI Pieter Abbeel, and Igor Mordatch. 2017. Multi-agent actor-critic for mixed cooperative-competitive environments. *Advances in neural information processing systems*, 30.

A Woodbury Max. 1950. Inverting modified matrices. In *Memorandum Rept. 42, Statistical Research Group*, page 4. Princeton Univ.

META. 2024. The llama 3 herd of models. *arXiv e-prints*, pages arXiv–2407.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. 2013. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.

Dmitry Moor. 2019. Data markets with dynamic arrival of buyers and sellers. In *Proceedings of the 14th Workshop on the Economics of Networks, Systems and Computation*, pages 1–6.

Xinyi Mou, Chen Qian, Wei Liu, Xuanjing Huang, and Zhongyu Wei. 2025. Ecolang: Efficient and effective agent communication language induction for social simulation. *arXiv preprint arXiv:2505.06904*.

OpenAI. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

Jian Pei. 2020. Data pricing–from economics to data science. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 3553–3554.

Jinghua Piao, Yuwei Yan, Jun Zhang, Nian Li, Junbo Yan, Xiaochong Lan, Zhihong Lu, Zhiheng Zheng, Jing Yi Wang, Di Zhou, and 1 others. 2025. Agentsociety: Large-scale simulation of llm-driven generative agents advances understanding of human behaviors and society. *arXiv preprint arXiv:2502.08691*.

Deepak Ravichandran and Nitish Korula. 2019. Effect of disabling third-party cookies on publisher revenue. *Google Report*.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Haoran Tang, Xueyao Sun, Shiqing Wu, Zhihong Cui, Guandong Xu, and Qing Li. 2025. Dybooster: Leveraging large language model as booster for dynamic recommendation. *Expert Systems with Applications*, 286:128080.

Jiachen T Wang and Ruoxi Jia. 2023. Data banzhaf: A robust data valuation framework for machine learning. In *International Conference on Artificial Intelligence and Statistics*, pages 6388–6421. PMLR.

Xiaoyang Wang, Yao Ma, Yiqi Wang, Wei Jin, Xin Wang, Jiliang Tang, Caiyan Jia, and Jian Yu. 2020. Traffic flow prediction via spatial temporal graph neural network. In *Proceedings of the web conference 2020*, pages 1082–1092.

Zengqing Wu, Run Peng, Shuyuan Zheng, Qianying Liu, Xu Han, Brian Inhyuk Kwon, Makoto Onizuka, Shaojie Tang, and Chuan Xiao. 2024. Shall we team up: Exploring spontaneous cooperation of competing llm agents. *arXiv preprint arXiv:2402.12327*.

Mingyan Xiao, Ming Li, and Jennifer Jie Zhang. 2023. Locally differentially private personal data markets using contextual dynamic pricing mechanism. *IEEE Transactions on Dependable and Secure Computing*, 20(6):5043–5055.

Chengxing Xie, Canyu Chen, Feiran Jia, Ziyu Ye, Shiyang Lai, Kai Shu, Jindong Gu, Adel Bibi, Ziniu

Hu, David Jurgens, and 1 others. 2024. Can large language model agents simulate human trust behavior? *Advances in neural information processing systems*, 37:15674–15729.

Aolin Xu and Maxim Raginsky. 2017. Information-theoretic analysis of generalization capability of learning algorithms. *Advances in neural information processing systems*, 30.

Yuwei Yan, Jinghua Piao, Xiaochong Lan, Chenyang Shao, Pan Hui, and Yong Li. 2025. Simulating generative social agents via theory-informed workflow design. *arXiv preprint arXiv:2508.08726*.

Shuhua Yang, Jiahao Zhang, Yilong Wang, Dongwon Lee, and Suhang Wang. 2026. Query-efficient agentic graph extraction attacks on graphrag systems. *arXiv preprint arXiv:2601.14662*.

Haifei Yu and Mengxiao Zhang. 2017. Data pricing strategy based on data quality. *Computers & Industrial Engineering*, 112:1–10.

Xu Yuan, Liangbo Ning, Wenqi Fan, and Qing Li. 2025. mkg-rag: Multimodal knowledge graph-enhanced rag for visual question answering. *arXiv preprint arXiv:2508.05318*.

YANG Yuzhe, Yifei Zhang, Minghao Wu, Kaidi Zhang, Yunmiao Zhang, Honghai Yu, Yan Hu, and Benyou Wang. 2025. Twinmarket: A scalable behavioral and social simulation for financial markets. In *ICLR 2025 Workshop on World Models: Understanding, Modelling and Scaling*.

Jiayao Zhang, Yuran Bi, Mengye Cheng, Jinfei Liu, Kui Ren, Qiheng Sun, Yihang Wu, Yang Cao, Raul Castro Fernandez, Haifeng Xu, and 1 others. 2024a. A survey on data markets. *arXiv preprint arXiv:2411.07267*.

Xinnong Zhang, Jiayu Lin, Libo Sun, Weihong Qi, Yihang Yang, Yue Chen, Hanjia Lyu, Xinyi Mou, Siming Chen, Jiebo Luo, and 1 others. 2024b. Electionsim: Massive population election simulation powered by large language model driven agents. *arXiv preprint arXiv:2410.20746*.

Zeyu Zhang, Jianxun Lian, Chen Ma, Yaning Qu, Ye Luo, Lei Wang, Rui Li, Xu Chen, Yankai Lin, Le Wu, and 1 others. 2024c. Trendsim: Simulating trending topics in social media under poisoning attacks with llm-based multi-agent system. *arXiv preprint arXiv:2412.12196*.

Zheyuan Zhang, Daniel Zhang-Li, Jifan Yu, Linlu Gong, Jinchang Zhou, Zhanxin Hao, Jianxiao Jiang, Jie Cao, Huiqin Liu, Zhiyuan Liu, and 1 others. 2025. Simulating classroom education with llm-empowered agents. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 10364–10379.

Yiding Zhu, Hongwei Zhang, Jiayao Zhang, Jinfei Liu, and Kui Ren. 2024. Dataprice: An interactive system for pricing datasets in data marketplaces. *Proceedings of the VLDB Endowment*, 17(12):4433–4436.

A Experimental Settings

The source code of EvoDM and Datamart-Agent is available at this GitHub repository.

A.1 Benchmark Construction Details

The benchmark is derived from the EvoDM introduced in Section 4, where each scenario admits a closed-form equilibrium and utility solution. Each scenario is generated by independently sampling user, platform, and buyer characteristics from predefined distributions. For each setting, we construct disjoint training, validation (with 200 examples for RL baselines), and test splits. In addition, we design OOD scenarios to evaluate robustness under distributional shifts and increased marketplace complexity. We summarize the benchmark construction for all experimental settings in Table 6.

Setting	Train #	Test #	α Range	β Range	#Platforms
CI	1000	200	$\mathbb{U}(0,5)$	$\mathbb{U}(0,1)$	2–6
Evolving	1000	200	$\mathbb{U}(0,5)$	$\mathbb{U}(0,1)$	2–8
β-INC	1000	200	$\mathbb{U}(0,5)$	$\mathbb{U}(0,1)$	2–6
CI (OOD)	–	200	$\mathbb{U}(0,10)$	$\mathbb{U}(0,2)$	6–8
β-INC (OOD)	–	200	$\mathbb{U}(0,10)$	$\mathbb{U}(0,2)$	6–8
Evolving (OOD)	–	200	$\mathbb{U}(0,10)$	$\mathbb{U}(0,2)$	6–16

Table 6: Benchmark construction settings.

In all settings, platform costs are sampled as $c_i \sim \mathbb{U}(0,1)$ and platform-user alignment parameters are sampled as $\gamma_i \sim \mathbb{U}(0,1)$. The OOD benchmarks jointly test the agent's ability to generalize across both *parameter shifts* (privacy sensitivity and buyer valuation) and *structural changes* (number of competing platforms), while maintaining consistency with analytically tractable equilibria.

A.2 Baselines

In our experiments, we compared our Datamart-Agent with the following RL- and LLM-driven baselines on EvoDM,

- **RL-ABM** (Lowe et al., 2017): RL-based ABM frameworks, in which each agent learns a policy from repeated environment interaction and then makes decisions.
- **LLM-State** (Zhang et al., 2024b): An LLM agent conditioned solely on its predefined role

and current state observations, without any memory of prior decisions.

- **LLM-Recent** (Zhang et al., 2024c; Piao et al., 2025): Extends the LLM-State by incorporating the k most recent decision trajectories as short-term flat memory.
- **LLM-I** (Li et al., 2024a): Incorporates the top-N most similar decision trajectories and prompts the agent to reason about how alternative actions affect others' decisions and utilities.
- **LLM-E** (Hua et al., 2024): Builds upon LLM-I by integrating both natural language and symbolic reasoning to infer how alternative actions affect others' decisions and utilities.

The LLM-driven baselines build on existing LLM-agent frameworks, adopting their design principles for multi-agent interaction and decision-making.

A.3 Metrics

We evaluate performance using two primary metrics. **1) Utility Gap** $\Delta\mathcal{U}$ depicts the deviation between predicted utilities $\hat{\mathcal{U}}_{*,j}$, $* \in \{u, i, b\}$ and ground-truth equilibrium utilities $\mathcal{U}_*$,

$$\Delta\mathcal{U} = \frac{1}{N}\sum_{j=1}^{N}\left|\hat{\mathcal{U}}_{*,j} - \mathcal{U}_*\right|. \qquad (13)$$

2) Pass@ε illustrates fraction of runs whose predicted utilities fall within an ε-tolerance of equilibrium,

$$\text{Pass@}\varepsilon = \frac{1}{N}\sum_{j=1}^{N}\mathbb{I}\big(|\hat{\mathcal{U}}_{*,j} - \mathcal{U}_*| \le \varepsilon\big), \qquad (14)$$

where $\mathbb{I}(\cdot)$ denotes the indicator function. For fine-grained evaluation, we additionally report accuracy for discrete actions (*i.e.*, $\boldsymbol{e}, \boldsymbol{a}$, and $\boldsymbol{b}$) and mean square error (MSE) of continuous decisions (*i.e.*, $\boldsymbol{\sigma}$ and $\boldsymbol{p}$).

A.4 Models and Hyperparameters

We instantiate the LLM-driven agents with three widely used instruction-tuned LLMs across different scales and families, including GPT-4O (OpenAI, 2023), LLAMA-3.1-8B-INSTRUCT (META, 2024), and QWEN-2.5-32B-INSTRUCT (Group, 2025). Unless otherwise noted, decoding is deterministic with `temperature=0` and `max_tokens=8192` to ensure stability and reproducibility. All other runtime hyperparameters follow the defaults of SGLANG[1].

For settings with training data (RQ1), we warm-up each LLM agent's memory by reviewing ground-truth equilibrium decisions from the training split without exposing test instances. For settings without training data (RQ2 and RQ3), we perform a 200-round self-play warm-up within the ABM environment. During simulation, the agent retrieves at most N_{mem}=10 prior game tree memory for our DATAMART-AGENT or flat decision traces for the baselines.

For RL-MAB baselines, we employ two representative RL models, DQN (Mnih et al., 2013), and PPO (Schulman et al., 2017), trained under identical conditions within the data marketplace ABM environment. All models adopt vectorized observations consistent with those used by the LLM-based Agents. The observation of RL-based agents includes decision-relevant historical variables, which already satisfies the Markov property for the policy learner. Therefore, we use an MLP policy model to capture states in data marketplace ABM environment and make actions. These RL models are optimized using the Adam optimizer with a learning rate of 3×10^{-4} and a discount factor of $\gamma = 0.99$. For DQN, we use a replay buffer of size 10^4, a batch size of 128, and an ϵ-greedy exploration strategy that decays from 0.9 to 0.02 over 10^3 steps. PPO follows standard configurations with generalized advantage estimation of $\lambda = 0.95$, a clipping ratio of 0.2, eight training epochs per policy update, a batch size of 64, a value loss coefficient of 0.5, and an entropy coefficient of 0.1.

B Extended Related Work

B.1 Data Marketplace Properties and Mechanisms

Research on data marketplaces is fundamentally shaped by the informational nature of data, which distinguishes it from traditional economic goods. In particular, data is commonly modeled as an *information good* whose value depends on signal quality and informativeness rather than exclusivity or consumption (Babaioff et al., 2012), naturally motivating information-theoretic abstractions for valuation and utility (Cover, 1999; Xu and Raginsky, 2017). These properties provide the conceptual foundation for formal models of data marketplaces.

[1] https://github.com/sgl-project/sglang

Early research on data marketplaces focuses on direct dataset trading, formalizing ownership, pricing, and access control (Yu and Zhang, 2017; Fernandez et al., 2020b; Zhang et al., 2024a). Subsequent work generalizes data products beyond static datasets to query-based markets (Koutris et al., 2015) and model-based markets (Chen et al., 2019; Jia et al., 2019; Liu et al., 2021), emphasizing arbitrage-free pricing, incentive compatibility, and fair revenue allocation under information reuse. These lines of work culminate in the three-layer data marketplace formulation (Fallah et al., 2024), which explicitly models strategic interactions among users, platforms, and buyers. Equilibrium analyzes under Stackelberg–Nash competition (Bi et al., 2024) and bargaining dynamics (Bi et al., 2025) further characterize incentive alignment under privacy and competition constraints.

A parallel literature examines regulation and competition in data exchange. Studies on data-sharing restrictions show that uniform policies may reduce welfare or innovation depending on market structure (Bimpikis et al., 2024; Argenziano and Bonatti, 2023; Ravichandran and Korula, 2019). Work on dynamic data sales and information design highlights how repeated interactions and data replicability complicate pricing and regulation under uncertainty (Abolhassani et al., 2017; Immorlica et al., 2021; Bergemann et al., 2018). Recent approaches propose adaptive mechanisms such as utility-balanced exchanges (Bhaskara et al., 2024) and performance-based pricing in federated settings (Li et al., 2024b).

Finally, learning-based approaches incorporate adaptive pricing and valuation into data marketplaces. These include mechanisms robust to strategic buyers (Castro Fernandez, 2022), task-agnostic data valuation methods (Amiri et al., 2023; Wang and Jia, 2023), and reinforcement-learning-driven market optimization (Li et al., 2021; Goktas, 2022). While effective under fixed assumptions, these models typically abstract away agent cognition and language-mediated interaction, limiting their applicability to more realistic and generalizable data marketplace analysis.

B.2 LLM-Based Agents in ABMs

Traditional ABMs rely on predefined rules or heuristics, producing emergent dynamics through repeated interactions (Gao et al., 2024). In contrast, recent LLM-based agents demonstrate strong capabilities in natural language reasoning, dialogue, and contextual adaptation (Jiang et al., 2025a,b; Yuan et al., 2025; Tang et al., 2025; Yang et al., 2026). They can efficiently decompose tasks hierarchically, plan over long horizons, and condition actions on memory, surpassing traditional agents in flexibility (Huang et al., 2022). This paradigm further exhibits negotiation and coordination behaviors through language, enabling multi-agent interaction without explicit protocol design (FAIR et al., 2022).

Empirical evaluations show that LLM agents reproduce human-like trust, reciprocity, cooperation, and risk preferences aligned with behavioral economics (Xie et al., 2024; Wu et al., 2024; Hu and Collier, 2024). Extensions incorporate motivational and educational theories to simulate adaptive learning and social dynamics (Yan et al., 2025; Zhang et al., 2025). More broadly, complex relational dependencies and dynamic interactions have also been studied in structured learning settings, including social recommendation and spatio-temporal forecasting (Fan et al., 2019; Wang et al., 2020). These agents have further been applied to economic and financial simulations, modeling heterogeneous decision-making and macro-level adaptation under policy interventions (Li et al., 2024a; Chen et al., 2025b; Yuzhe et al., 2025).

Despite these advances, a key challenge remains in ensuring equilibrium-consistent decision-making by LLM agents in strategic environments. Most existing studies emphasize behavioral realism or emergent social patterns, whereas our work focuses on enhancing LLM agents' decision rationality in data marketplace simulations, thereby enabling principled analysis of strategic interaction, incentive alignment, and regulatory feedback.

C Theorem Proofs and Examples

In this section, we provide full derivations, proofs, and illustrative examples for the equilibria and utilities used as ground truth in Section 3. We organize the materials by (i) static CI equilibrium foundations, (ii) temporally evolving entry, and (iii) buyer-valuation incomplete information (β-INC).

C.1 Preliminary (Static CI)

This section presents the formal propositions and proofs that underpin the equilibrium structure of the static, complete-information three-layer data marketplace introduced in Section 3.1. Under the standard Gaussian assumptions adopted in Fallah et al. (2024), we characterize how platform signals induce information revelation, how buyer-facing noise affects information aggregation, and how these primitives jointly determine privacy leakage, equilibrium pricing, and agent utilities.

Before presenting the technical results, we briefly explain the economic and informational rationale behind the utility functions defined in Definition 2. Each agent's utility is expressed in terms of mutual information to directly capture the value and cost of information flows in the marketplace. Specifically, platform i's utility $\mathcal{U}_i(\boldsymbol{\sigma}, \boldsymbol{e}, \boldsymbol{a}, \boldsymbol{p}, \boldsymbol{b})$ consists of a service component proportional to the user-side revealed information $\mathcal{I}_i = \mathcal{I}(x_i^\top \theta \mid s_i)$ whenever the user shares data ($a_i = 1$), together with revenue $b_i p_i$ from selling noisy signals to the buyer, minus a fixed entry cost c_i. The user's utility $\mathcal{U}_u(\boldsymbol{\sigma}, \boldsymbol{e}, \boldsymbol{a}, \boldsymbol{p}, \boldsymbol{b})$ trades off service gains $\sum_i a_i e_i \mathcal{I}_i$ against a privacy loss measured by the total information leaked to the buyer, $\mathcal{I}(\boldsymbol{y}^\top \theta \mid \tilde{\boldsymbol{s}})$. The buyer's utility $\mathcal{U}_b(\boldsymbol{\sigma}, \boldsymbol{e}, \boldsymbol{a}, \boldsymbol{p}, \boldsymbol{b})$ is linear in the aggregated information obtained from all purchased platforms, scaled by valuation parameter β, and net of total payments.

Formulating utilities directly in terms of mutual information ensures that noise choices, information substitutability across platforms, and privacy–utility trade-offs are explicitly and tractably represented, enabling closed-form equilibrium characterization under Gaussian assumptions.

Proposition A1 (User-Side Revealed Information)**.** *For data signal $s_i = x_i^\top \theta + \varepsilon_i$ shared to platform i with $\varepsilon_i \sim \mathcal{N}(0, 1)$ and $\|x_i\| = 1$, the revealed information is,*

$$\mathcal{I}_i = \frac{1}{2} \log \left(1 + \frac{\mathrm{Var}(x_i^\top \theta)}{\mathrm{Var}(\varepsilon_i)}\right) = \frac{1}{2}. \tag{15}$$

Proof. Since $\theta \sim \mathcal{N}(0, I_d)$ and $\|x_i\| = 1$, we have $x_i^\top \theta \sim \mathcal{N}(0, 1)$ and $\mathrm{Var}(x_i^\top \theta) = 1$. The signal $s_i = x_i^\top \theta + \varepsilon_i$ is a scalar Gaussian channel with independent noise $\varepsilon_i \sim \mathcal{N}(0, 1)$. Thus, $\mathcal{I}_i = I(x_i^\top \theta; s_i) = \frac{1}{2} \log\left(1 + \frac{\mathrm{Var}(x_i^\top \theta)}{\mathrm{Var}(\varepsilon_i)}\right) = \frac{1}{2} \log(2) = \frac{1}{2}$. □

Proposition A2 (Buyer-Side Received Information)**.** *Let $S = \{i \mid e_i = a_i = b_i = 1\} \subseteq [K]$ be the set of platforms whose data reach the buyer, $\boldsymbol{m}_S = [\gamma_i]_{i \in S}$, and*

$$(\boldsymbol{M}_S)_{ij} = \begin{cases} 2 + \sigma_i^2, & i = j, \\ \gamma_i \gamma_j, & i \neq j. \end{cases} \tag{16}$$

Then, $\mathcal{I}(\boldsymbol{\sigma}_S) = \boldsymbol{m}_S^\top \boldsymbol{M}_S^{-1} \boldsymbol{m}_S$.

Proof. Let $\tilde{\boldsymbol{s}}_S = [\tilde{s}_i]_{i \in S}$ denote the vector of buyer-observed signals. Under the linear-Gaussian model in Section 3, $\tilde{\boldsymbol{s}}_S$ is jointly Gaussian with the target scalar $y^\top \theta$ and admits the standard mutual-information form $I(y^\top \theta; \tilde{\boldsymbol{s}}_S) = \frac{1}{2} \log\left(\frac{\mathrm{Var}(y^\top \theta)}{\mathrm{Var}(y^\top \theta \mid \tilde{\boldsymbol{s}}_S)}\right)$. Writing the conditional variance via the Schur complement yields the closed form $\mathcal{I}(\boldsymbol{\sigma}_S) = \boldsymbol{m}_S^\top \boldsymbol{M}_S^{-1} \boldsymbol{m}_S$, with $\boldsymbol{m}_S$ and $\boldsymbol{M}_S$ defined in Proposition A2. □

Corollary A1 (Individual Platform Received Information)**.** *For $|S| = 1$, Proposition A2 yields,*

$$\mathcal{I}(\sigma_i) = \frac{\gamma_i^2}{2 + \sigma_i^2}. \tag{17}$$

Proof. When $|S| = 1$, $\boldsymbol{m}_S = [\gamma_i]$ and $\boldsymbol{M}_S = [2 + \sigma_i^2]$. Therefore $\boldsymbol{m}_S^\top \boldsymbol{M}_S^{-1} \boldsymbol{m}_S = \gamma_i^2 / (2 + \sigma_i^2)$, proving the claim. □

Proposition A3 (Leaked Information and Equal-Leakage Structure)**.** *Fix a participating platform set S, and assume the user shares with every entered platform. An equilibrium noise profile exists iff $\alpha > \frac{K-1}{2}$ and is characterized by the common ratio,*

$$\frac{2 + \sigma_i^2}{\gamma_i^2} = 2\alpha - K + 1. \tag{18}$$

The buyer's total information equals the user's privacy boundary,

$$\mathcal{I}(\boldsymbol{\sigma}_S) = \frac{K}{2\alpha}. \tag{19}$$

Proof (sketch). (i) By Proposition A2, the buyer-side leakage $\mathcal{I}(\boldsymbol{\sigma}_S)$ is a smooth function of $\{\sigma_i\}_{i \in S}$. (ii) When the user shares with all entered platforms, the user's best response makes the privacy term binding, which implies $\mathcal{I}(\boldsymbol{\sigma}_S) = K/(2\alpha)$. (iii) Platforms' optimality conditions for maximizing revenue, subject to the binding leakage constraint, force the equal-ratio structure $\frac{2+\sigma_i^2}{\gamma_i^2} = 2\alpha - K + 1$ for all $i \in S$. Feasibility requires $2\alpha - K + 1 > 0$, i.e., $\alpha > (K - 1)/2$. □

Proposition A4 (Platform Pricing Equilibrium (Marginal Contribution)). *Given entry $\boldsymbol{e}$, sharing $\boldsymbol{a}$, and noise $\boldsymbol{\sigma}$, the optimal price posted by platform i is its marginal contribution to the buyer:*

$$p_i^* = \beta\left(\mathcal{I}(\boldsymbol{\sigma}_S) - \mathcal{I}\left(\boldsymbol{\sigma}_{S\setminus\{i\}}\right)\right). \tag{20}$$

Proof. Fix entry $\boldsymbol{e}$, sharing $\boldsymbol{a}$, and noise $\boldsymbol{\sigma}$. Given $\boldsymbol{\sigma}$, the buyer's utility is linear in purchased information, and the buyer selects S by comparing per-unit value with induced per-unit prices. In equilibrium, each platform sets price equal to its marginal information contribution to the buyer, yielding Eq. (20). □

These propositions and corollaries characterize the subgame-perfect equilibrium of the three-layer marketplace and serve as ground truth for the multi-agent simulation.

C.2 Temporally Evolving Marketplace

In this section, we formalize equilibrium behavior in a temporally evolving data marketplace with sequential platform entry. We provide proofs for the entrant noise-price characterization and the resulting per-round utility decomposition, as stated in the main paper.

Proposition A5 establishes the equilibrium noise-price choice of a platform entering in round t, formally deriving the closed-form expressions of temporally evolving marketplaces under a binding marginal privacy constraint.

Proposition A5 (Noise–Price Pair for Round-t Entrants). *Fix any round $t \geq 2$ with incumbents S_{t-1}. For each new entrant $j \notin S_{t-1}$, the equilibrium noise level satisfies,*

$$\sigma_j^2(t) = 2\alpha\gamma_j^2\left(1 - \mathcal{I}(\boldsymbol{\sigma}_{S_{t-1}})\right)^2 + \gamma_j^2\mathcal{I}(\boldsymbol{\sigma}_{S_{t-1}}) - 2, \tag{21}$$

whenever the right-hand side is non-negative; otherwise, the entrant cannot be admitted without violating the privacy boundary. The corresponding equilibrium entry-time price is,

$$p_j^{(t)} = \frac{\beta_t}{2\alpha}. \tag{22}$$

Proof. At round t, a new platform $j \in S_t^{\text{New}}$ can be accepted by the user only if the incremental buyer-side information $\Delta\mathcal{I}_j^{(t)} := \mathcal{I}(\boldsymbol{\sigma}_{S_{t-1}} \cup \{\sigma_j\}) - \mathcal{I}(\boldsymbol{\sigma}_{S_{t-1}})$ satisfies the marginal privacy constraint $\frac{1}{2\alpha}$, which arises from the condition that the incremental service gain $\mathcal{I}_j = \frac{1}{2}$ equals the privacy cost $\alpha\Delta\mathcal{I}_j^{(t)}$ according to Eq. (2). Under the Gaussian assumption and the substitute-good information structure in Proposition A2, this marginal gain can be expressed using the Woodbury identity (Max, 1950) as,

$$\Delta\mathcal{I}_j^{(t)} = \frac{\gamma_j^2\left(1 - \mathcal{I}(\boldsymbol{\sigma}_{S_{t-1}})\right)^2}{2 + \sigma_j^2 - \gamma_j^2\mathcal{I}(\boldsymbol{\sigma}_{S_{t-1}})}.$$

Solving for σ_j^2 gives the unique candidate noise in Eq. (21). Feasibility requires $\sigma_j^2(t) \geq 0$; otherwise, no valid noise level satisfies the privacy constraint and the entrant is excluded. If feasible and accepted, the entrant's marginal information contribution equals $\frac{1}{2\alpha}$, implying an equilibrium price proportional to the buyer's valuation, yielding Eq. (22) by Proposition A4. □

Proposition 1 verifies the per-round utility decomposition, showing that user surplus is driven to zero, incumbents earn constant flow payoffs, entrants receive entry-time revenues, and buyer utility evolves with cumulative prices and information.

Proof. The user utility $U_u^{(t)}$ follows from Proposition A3, which implies the binding leakage structure and hence $\mathcal{U}_u^{(t)} = 0$. Incumbents keep (σ_i, p_i) and obtain constant per-round utility in Eq. (5). Entrants satisfy Eq. (22), hence their per-round utility is Eq. (6). Buyer $\boldsymbol{y}_t$ earns $\beta_t\mathcal{I}(\boldsymbol{\sigma}_t)$ and pays incumbents P_{t-1} plus m_t entrants at price $\beta_t/(2\alpha)$, giving Eq. (7). □

Example A1 (Three-Round Evolution with Sequential Entry). *In the first round, let $\alpha = 4$, $\beta_1 = 1$, and incumbent platforms $(\gamma_1, \gamma_2) = (1.0, 0.8)$ with costs $(c_1, c_2) = (0.55, 0.40)$. By Proposition A3, $\sigma_1^2 = 5.00$, $\sigma_2^2 = 2.48$, and $\mathcal{I}(\boldsymbol{\sigma}_{S_1}) = \frac{1}{4}$. Prices from Proposition A4 yield $p_1^{(1)} = p_2^{(1)} = \frac{3}{28} = 0.107$, and the corresponding utilities are $\mathcal{U}_u^{(1)} = 0$, $\mathcal{U}_1^{(1)} = 0.057$, $\mathcal{U}_2^{(1)} = 0.207$, $\mathcal{U}_b^{(1)} = 0.036$.*

In the second round, a new buyer with $\beta_2 = 1$ arrives, and the pre-entry information leakage is $\mathcal{I}(\boldsymbol{\sigma}_{S_1}) = \frac{1}{4}$. A new platform $j = 3$ with $(\gamma_3, c_3) = (1.0, 0.30)$ enters. By Proposition A5, $\sigma_3^2 = 2.75$ and $p_3^{(2)} = 0.125$. The updated information leakage is $\mathcal{I}(\boldsymbol{\sigma}_{S_2}) = 0.375$, and utilities are $\mathcal{U}_u^{(2)} = 0$, $\mathcal{U}_1^{(2)} = 0.057$, $\mathcal{U}_2^{(2)} = 0.207$, $\mathcal{U}_3^{(2)} = 0.325$, $\mathcal{U}_b^{(2)} = 0.036$.

In the third round, buyer $\beta_3 = 1.2$ arrives and a new platform $j = 4$ with $(\gamma_4, c_4) = (0.8, 0.35)$

enters. By Proposition A5, $\sigma_4^2 = 0.24$ and $p_4^{(3)} = 0.15$. This leads to $\mathcal{I}(\boldsymbol{\sigma}_{S_3}) = 0.5$. Per-round utilities are $\mathcal{U}_u^{(3)} = 0$, $\mathcal{U}_1^{(3)} = 0.057$, $\mathcal{U}_2^{(3)} = 0.207$, $\mathcal{U}_3^{(3)} = 0.325$, $\mathcal{U}_4^{(3)} = 0.300$, and buyer utility $\mathcal{U}_b^{(3)} = 0.111$.

C.3 β-INC Marketplace

In this section, we provide the formal equilibrium characterizations and illustrative examples for the data marketplace under buyer-valuation incomplete information, including optimal per-unit pricing, user sharing cutoffs, equilibrium noise design, and utility outcomes in low- and high-privacy regimes.

The optimal per-unit price defined in Section 3.3 can be obtained by,

Lemma A1 (Optimal Per-Unit Price)**.** *Assume $\beta \sim \mathbb{U}[\underline{\beta}, \overline{\beta}]$. Given any $\mathcal{I}(\sigma_i)$ and the expected revenue of platform i, $R_i(z_i) = z_i \mathcal{I}(\sigma_i)[1 - F(z_i)]$, the revenue is maximized at $z^* = \max\left\{\underline{\beta}, \frac{\overline{\beta}}{2}\right\}$.*

Proof. Under the uniform prior $F(z_i) = \frac{z_i - \underline{\beta}}{\overline{\beta} - \underline{\beta}}$, we have $R_i(z_i) \propto z_i(\overline{\beta} - z_i)$, a concave quadratic maximized at $z_i = \overline{\beta}/2$. Projecting back to $[\underline{\beta}, \overline{\beta}]$ yields $z^* = \max\{\underline{\beta}, \overline{\beta}/2\}$. □

Lemma A2 (User Cut-off in Buyer INC)**.** *The user shares its data with all entered K platforms iff,*

$$\alpha \leq \alpha_{\text{crit}} := \frac{K}{2\Pr(\beta \geq z^*)\mathcal{I}(\boldsymbol{\sigma}_S)}. \tag{23}$$

Proof. If the user shares with all entered platforms, Eq. (2) and Proposition A1 yield

$$\mathbb{E}_\beta[\mathcal{U}_u] = \frac{K}{2} - \alpha \Pr(\beta \geq z^*)\mathcal{I}(\boldsymbol{\sigma}_S).$$

If the user shares with no platform, the utility is 0. Therefore, sharing all is preferred iff $\mathbb{E}_\beta[\mathcal{U}_u] \geq 0$, which yields the stated condition. □

C.3.1 Low-privacy regime.

When $\alpha \leq \alpha_{\text{crit}}$, the user is willing to share even with zero noise. Each platform maximizes expected revenue by choosing $\sigma_i^* = 0$ and posting $p_i^* = z^* \mathcal{I}(\sigma_i^*) = \frac{z^* \gamma_i^2}{2}$.

Proposition A6 (Equilibrium at Low-Privacy Regime)**.** *If $\alpha \leq \alpha_{\text{crit}}$, any sub-game-perfect equilibrium satisfies $\sigma_i^* = 0$ and $p_i^* = z^* \frac{\gamma_i^2}{2}$ for all $i \in S$. The resulting expected utilities are,*

$$\mathbb{E}_\beta[\mathcal{U}_u] = \frac{K}{2} - \alpha \Pr(\beta \geq z^*)\mathcal{I}(\boldsymbol{\sigma}_S^*), \tag{24}$$

$$\mathbb{E}_\beta[\mathcal{U}_i] = \frac{1}{2} - c_i + \Pr(\beta \geq z^*) p_i^*, \tag{25}$$

$$\mathbb{E}_\beta[\mathcal{U}_b] = \Pr(\beta \geq z^*)(\mathbb{E}[\beta \mid \beta \geq z^*] - z^*)\mathcal{I}(\boldsymbol{\sigma}_S^*). \tag{26}$$

Proof. Given $\alpha \leq \alpha_{\text{crit}}$, Lemma A2 implies the user shares with all entered platforms even when $\sigma_i = 0$. Platforms then maximize expected revenue by setting per-unit price z^* (Lemma A1) and hence $p_i^* = z^* \mathcal{I}(\sigma_i^*) = z^* \gamma_i^2 / 2$. The expected utilities follow by substituting σ_i^*, p_i^* into Eqs. (1)–(3) and taking expectations over β. □

C.3.2 High-privacy regime.

When $\alpha > \alpha_{\text{crit}}$, the user shares only if the *expected* leakage satisfies,

$$\Pr(\beta \geq z^*) \cdot \mathcal{I}(\boldsymbol{\sigma}_S) = \frac{K}{2\alpha}. \tag{27}$$

Platforms then choose noise and price to maximize expected revenue subject to this binding constraint.

Theorem A1 (Optimal Noise and Price at High-Privacy Regime)**.** *In the high-privacy regime $\alpha > \alpha_{\text{crit}}$, platforms choose,*

$$\sigma_i^{*2} = \gamma_i^2 \left(2\alpha \Pr(\beta \geq z^*) - K + 1\right) - 2, \tag{28}$$

$$p_i^* = \frac{z^*}{2\alpha \Pr(\beta \geq z^*) - K + 1}, \tag{29}$$

which satisfy Eq. (27) with minimal noise while maximizing expected revenue.

Proof. Let $\phi_i := \frac{\gamma_i^2}{2 + \sigma_i^2 - \gamma_i^2} > 0$. Woodbury formula gives, $\mathcal{I}(\boldsymbol{\sigma}_S) = \frac{\sum_i \phi_i}{1 + \sum_i \phi_i}$, and Eq. (27) becomes,

$$\sum_{i \in S} \phi_i = \frac{K}{2\alpha \Pr(\beta \geq z^*)} - 1.$$

The expected revenue of platform i is,

$$R_i(\phi_i) = \Pr(\beta \geq z^*) z^* \frac{\phi_i}{1 + \phi_i}.$$

Up to the positive constant $\Pr(\beta \geq z^*) z^*$, the objective is $f(\phi_i) := \frac{\phi_i}{1 + \phi_i}$, which is strictly concave since $f''(\phi_i) = -\frac{2}{(1+\phi_i)^3} < 0$. By first-order conditions under a linear coupling constraint, the optimum equalizes ϕ_i across i:

$$\phi_i = \frac{1}{K}\Big(\frac{K}{2\alpha\Pr(\beta\geq z^*)}-1\Big) = \frac{1}{2\alpha\Pr(\beta\geq z^*)-K}.$$

Substituting back yields,

$$\sigma_i^{*2} = \gamma_i^2\left(2\alpha\Pr(\beta\geq z^*)-K+1\right)-2.$$

Finally, using $p_i^* = z^*\mathcal{I}(\sigma_i^*)$ and $\mathcal{I}(\sigma_i^*) = \gamma_i^2/(2+\sigma_i^{*2})$ gives,

$$p_i^* = \frac{z^*}{2\alpha\Pr(\beta\geq z^*)-K+1}.$$

□

Proposition A7 (Utilities in the High-Privacy Regime)**.** *Assume $\alpha > \alpha_{\text{crit}}$ and the optimal noise–price pair of Theorem A1. In equilibrium, $e_i = a_i = 1$ for all $i \in S$, and $b_i = 1$ only when $\beta \geq z^*$. Then the expected utilities are,*

$$\mathbb{E}_\beta[\mathcal{U}_u]=0, \tag{30}$$

$$\mathbb{E}_\beta[\mathcal{U}_i]=\frac{1}{2}-c_i+\frac{\Pr(\beta\geq z^*)z^*}{2\alpha\Pr(\beta\geq z^*)-K+1}, \tag{31}$$

$$\mathbb{E}_\beta[\mathcal{U}_b]=\frac{K}{2\alpha}\mathbb{E}[\beta\,|\,\beta\geq z^*]-\frac{K\Pr(\beta\geq z^*)z^*}{2\alpha\Pr(\beta\geq z^*)-K+1}. \tag{32}$$

Proof. For the user, Proposition A1 yields a total service gain of $K/2$. In the high-privacy regime, the expected leakage constraint binds: $\Pr(\beta \geq z^*)\mathcal{I}(\boldsymbol{\sigma}_S^*) = K/(2\alpha)$, therefore,

$$\mathbb{E}_\beta[\mathcal{U}_u] = \frac{K}{2} - \alpha\cdot\frac{K}{2\alpha} = 0.$$

For platform i, the platform receives p_i^* only when $\beta \geq z^*$, hence $\mathbb{E}_\beta[\mathcal{U}_i] = \frac{1}{2} - c_i + \Pr(\beta \geq z^*)p_i^*$; substituting p_i^* from Theorem A1 yields the stated form. For the buyer, conditional on $\beta \geq z^*$, utility is $\beta\mathcal{I}(\boldsymbol{\sigma}_S^*) - \sum_{i\in S} p_i^*$, otherwise 0. Taking expectations and substituting the binding leakage and p_i^* yields the expression. □

Example A2 (Low-Privacy Regime)**.** *Let $\gamma = (1.0, 0.8, 0.6)$, $c = (0.55, 0.40, 0.30)$, and $\alpha = 1.5$. Assume $\beta \sim \mathbb{U}[0, 1]$, so $z^* = \overline{\beta}/2 = 0.5$ and $\Pr(\beta \geq z^*) = 0.5$. In the low-privacy regime, $\sigma_i^* = 0$ and $p_i^* = z^* \cdot \frac{\gamma_i^2}{2} = (0.25, 0.16, 0.09)$. Using Proposition A6, platform utilities are $\mathbb{E}_\beta[\mathcal{U}_1] = 0.075, \mathbb{E}_\beta[\mathcal{U}_2] = 0.180, \mathbb{E}_\beta[\mathcal{U}_3] = 0.245$. The user's utility is $\mathbb{E}_\beta[\mathcal{U}_u] = 1.029$ and the buyer utility is $\mathbb{E}_\beta[\mathcal{U}_b] = 0.079$.*

Example A3 (High-Privacy Regime)**.** *Let $\beta \sim \mathbb{U}[0, 1]$, so $z^* = 0.5$ and $\Pr(\beta \geq z^*) = 0.5$. Take $\gamma = (1.0, 0.8, 0.6)$, entry costs $c = (0.5, 0.40, 0.30)$, and choose $\alpha = 8 > \alpha_{\text{crit}}$ so that the market operates in the high-privacy regime. With $K = 3$ and $2\alpha\Pr(\beta \geq z^*) - K + 1 = 6$, Theorem A1 gives $\sigma^{*2} = (4.00, 1.84, 0.16)$ and $p_i^* = 0.0833$. Using Proposition A7, $\mathbb{E}_\beta[\mathcal{U}_u] = 0$, $\mathbb{E}_\beta[\mathcal{U}_1] = 0.042$, $\mathbb{E}_\beta[\mathcal{U}_2] = 0.142$, $\mathbb{E}_\beta[\mathcal{U}_3] = 0.242$, and $\mathbb{E}_\beta[\mathcal{U}_b] = 0.016$.*

C.4 Justification for Non-Exclusive Data Trading

This section justifies two modeling choices, similar to those in the general data marketplace studies, made in Sections 3 and 4: (i) modeling a single buyer arrival in each round, and (ii) omitting explicit buyer competition in the data trading stage. We demonstrate that both choices are without loss of generality under non-exclusive data trading and information-based pricing.

We fix a round t and let $S \subseteq [K]$ denote the set of platforms whose data reach the buyer. Let $\sigma_S = \{\sigma_i\}_{i\in S}$ be the corresponding noise profile, and let $I(\sigma_S)$ denote the buyer-side received information as defined in Prop. A2. Platform prices $\{p_i\}_{i\in S}$ are determined by marginal information contribution as Prop. A4. For any buyer with valuation $\beta \geq 0$, the buyer utility is

$$U_b = \beta I(\sigma_S) - \sum_{i\in S} b_i p_i, \tag{33}$$

where $b_i \in \{0, 1\}$ indicates whether the buyer purchases from platform i.

C.4.1 Equivalence of Single-Buyer and Multi-Buyer Formulations

We first demonstrate that explicitly modeling multiple buyers within the same round is equivalent to a single-buyer formulation with an aggregated valuation.

Theorem A2 (Buyer Aggregation Equivalence)**.** *Suppose that in round t there are N_t buyers indexed by $k = 1, \ldots, N_t$, each with valuation $\beta_{t,k}$, and buyers purchase data independently under non-exclusive access. Then, for any platform set S and noise profile σ_S, equilibrium platform pricing and platform incentives are equivalent to a* *single-buyer formulation with valuation,*

$$\beta_t^{\mathrm{agg}} := \sum_{k=1}^{N_t} \beta_{t,k}. \tag{34}$$

Proof. Each buyer k derives utility,

$$U_{b,k} = \beta_{t,k} I(\sigma_S) - \sum_{i \in S} b_{i,k} p_i. \tag{35}$$

Because data access is non-exclusive, a buyer's purchase does not affect the information received by other buyers, nor does it constrain platforms from serving additional buyers. Hence, buyer utilities are additively separable. Summing over all buyers yields,

$$\sum_{k=1}^{N_t} U_{b,k} = \left(\sum_{k=1}^{N_t} \beta_{t,k} \right) I(\sigma_S) - \sum_{i \in S} \left(\sum_{k=1}^{N_t} b_{i,k} \right) p_i. \tag{36}$$

By Prop A4, each platform's optimal price is proportional to its marginal information contribution $I(\sigma_S) - I(\sigma_{S \setminus \{i\}})$, which is independent of buyer identity. Therefore, the only buyer-side quantity affecting pricing is the aggregated valuation $\sum_k \beta_{t,k}$, implying the stated equivalence. □

The aggregation result extends naturally to stochastic buyer arrivals.

Corollary A2 (Temporal Distributional Equivalence)**.** *If buyers' valuations are independently drawn from a common distribution F, then modeling a single buyer with valuation $\beta_t \sim F$ in each round yields the same expected pricing and platform behavior by explicitly modeling multiple buyers per round.*

Proof. By Theorem A2, pricing depends only on the aggregated valuation $\sum_k \beta_{t,k}$. Under i.i.d. draws,

$$\mathbb{E}\left[\sum_{k=1}^{N_t} \beta_{t,k} \right] = N_t \mathbb{E}[\beta]. \tag{37}$$

An equivalent expected valuation mass is generated by sequential single-buyer arrivals with $\beta_t \sim F$ over time. Since platform pricing and noise choices depend only on the realized or expected valuation, both formulations induce the same equilibrium behavior. □

C.4.2 Irrelevance of Explicit Buyer Competition

We now show that explicit multiple-buyer competition does not affect equilibrium outcomes in the data trading stage.

Proposition A8 (Irrelevance of Buyer Competition)**.** *In the data marketplace, explicit strategic interaction among buyers does not affect equilibrium platform pricing or platform incentives, provided that buyer competition does not alter access feasibility or the information term $I(\sigma_S)$.*

Proof. Under non-exclusive access, the buyer-side received information $I(\sigma_S)$ depends only on the platform set S and the noise profile σ_S, and is independent of the number of buyers purchasing the data. Platform pricing is determined by marginal information contribution as Prop. A4, which depends on $I(\sigma_S)$ and buyer valuation β.

Buyer competition occurs downstream after data acquisition and affects buyers' realized payoffs in their respective tasks or markets. Its effect on the data trading stage is fully captured by buyers' willingness-to-pay, represented by β or its distribution. Therefore, modeling explicit buyer competition introduces no additional pricing primitives unless exclusivity, capacity constraints, or buyer-side externalities are imposed. □

Remark A1 (Exclusive Data Trading)**.** *The above equivalence fails if data access is exclusive, capacity-constrained, or if buyer-side congestion affects $I(\sigma_S)$, which is not a general characteristic of the data marketplace mechanism. In such cases, data becomes effectively rivalrous, and the market reduces to a setting with scarce goods or auctions, which is outside the scope of this work.*

D Prompts

In this section, we illustrate the prompts of role-conditioned perception, dynamic game tree memory, and mechanism-guided reflection decision of each role in the EvoDM.

Platform Agent Prompts

System Prompt:

You are an agent in a three-layer data marketplace. You need to make rational decisions based on the current market situation.

Dynamic Role-Conditioned Perception.

- You are a platform agent at the [stage] stage.
- [platform_role_description_for_][stage]

Stage description.

- [platform_stage_description_for_][stage]

Dynamic Game Tree Memory. Past episodes are stored as a game tree. Each root node corresponds to one environment with parameters [env_params]. Each episode is stored as a root-to-leaf path formatted as:

- Root: [environment_description]
- [stage_1] stage decision: [decision_description_1]
- [stage_2] stage decision: [decision_description_2]
- ...
- Leaves: [utility_description including U_user, U_platforms, U_buyer]

When current environment parameters [current_env_params] are available, select roots whose [env_params] are sufficiently similar; if none are similar, use a small number of the most recent roots. Provide up to [max_paths] such paths as [memory_context].

Mechanism-Guided Explicit Reasoning. At stage [stage], reason explicitly as a platform using the stage-specific mechanism.

1. Utility function. Use platform utility:

- U_i = [a_i (I_i + b_i p_i) - c_i] * e_i

2. Meaning of symbols. Briefly explain the symbols relevant to the current platform stage, especially e_i, a_i, I_i, b_i, p_i, c_i, and when relevant sigma_i, gamma_i, alpha, beta.

3. Analytical impact. Using the formula, explain how changing the platform decision variable [decision_variable_for_platform_][stage] affects platform utility and, when relevant, user participation and buyer demand.

4. Stepwise impact. Given the current parameter values, identify which action maximizes U_i at the [stage] stage.

Final decision rule. Base your final decision on both: (a) patterns from [memory_context], and (b) the formal mechanism above.

User Prompt:

Previous Experiences: [memory_context]

Current Market Situation: [observation_text]

Decision Required: [platform_stage_decision_question]

Return: [stage_response_format]

Please respond only with a JSON object with the following fields:

- "reasoning": your decision-making process and key considerations.
- "action": your final decision value.

No other text or comments are allowed.

User Agent Prompts

System Prompt:

You are an agent in a three-layer data marketplace. You need to make rational decisions based on the current market situation.

Dynamic Role-Conditioned Perception.

- You are a data owner (user) at the user sharing stage.
- Decide whether to share your data with different platforms.
- Consider privacy concerns and potential benefits from sharing.

Stage description.

- This is the user sharing stage.
- Consider privacy costs and potential benefits from data sharing.

Dynamic Game Tree Memory. Past episodes are stored as a game tree. Each root node corresponds to one environment with parameters [env_params]. Each episode is stored as a root-to-leaf path formatted as:

- Root: [environment_description]
- [stage_1] stage decision: [decision_description_1]
- [stage_2] stage decision: [decision_description_2]
- ...
- Leaves: [utility_description including U_user, U_platforms, U_buyer]

When current environment parameters [current_env_params] are available, select roots whose [env_params] are sufficiently similar; if none are similar, use a small number of the most recent roots. Provide up to [max_paths] such paths as [memory_context].

Mechanism-Guided Explicit Reasoning. Reason explicitly as a user.

1. Utility function. Use user utility:

- U_u = \sum _i a_i e_i I_i - alpha * I_total

2. Meaning of symbols. Briefly explain a_i, e_i, I_i, alpha, and I_total.

3. Analytical impact. Using the formula, explain how the sharing decision a_i changes user utility, emphasizing the trade-off between service benefit and privacy loss.

4. Stepwise impact. Given the current parameter values, identify which sharing action maximizes U_u.

Final decision rule. Base your final decision on both: (a) patterns from [memory_context], and (b) the formal mechanism above.

User Prompt:

Previous Experiences: [memory_context]

Current Market Situation: [observation_text]

Decision Required: Which platforms should you share your data with?

Return: [stage_response_format]

Please respond only with a JSON object with the following fields:

- "reasoning": your decision-making process and key considerations.
- "action": your final decision value.

No other text or comments are allowed.

Buyer Agent Prompts

System Prompt:

You are an agent in a three-layer data marketplace. You need to make rational decisions based on the current market situation.

Dynamic Role-Conditioned Perception.

- You are a data buyer at the buyer selection stage.
- Decide which platform's data to purchase based on quality, price, and your needs.

Stage description.

- This is the buyer selection stage.
- Consider data quality, prices, and your valuation.

Dynamic Game Tree Memory. Past episodes are stored as a game tree. Each root node corresponds to one environment with parameters [env_params]. Each episode is stored as a root-to-leaf path formatted as:

- Root: [environment_description]
- [stage_1] stage decision: [decision_description_1]
- [stage_2] stage decision: [decision_description_2]
- ...
- Leaves: [utility_description including U_user, U_platforms, U_buyer]

When current environment parameters [current_env_params] are available, select roots whose [env_params] are sufficiently similar; if none are similar, use a small number of the most recent roots. Provide up to [max_paths] such paths as [memory_context].

Mechanism-Guided Explicit Reasoning. Reason explicitly as a buyer.

1. Utility function. Use buyer utility:

- U_b = beta * I(sigma, e, a, b) - \sum _i b_i p_i

2. Meaning of symbols. Briefly explain beta, I(sigma,e,a,b), b_i, and p_i.

3. Analytical impact. Using the formula, explain how the purchase decision b_i affects buyer utility, emphasizing the trade-off between information gain and payment cost.

4. Stepwise impact. Given the current parameter values, identify which purchase action maximizes U_b.

Final decision rule. Base your final decision on both: (a) patterns from [memory_context], and (b) the formal mechanism above.

User Prompt:

Previous Experiences: [memory_context]

Current Market Situation: [observation_text]

Decision Required: Which platform's data should you purchase?

Return: [stage_response_format]

Please respond only with a JSON object with the following fields:

- "reasoning": your decision-making process and key considerations.
- "action": your final decision value.

No other text or comments are allowed.