MECHANISM PARAMETER OPTIMIZATION UNDER BEHAVIORAL UNCERTAINTY VIA SEQUENTIAL EXPERIMENTAL DESIGN

Anonymous authors
Paper under double-blind review

ABSTRACT

In strategic environments, mechanism parameter optimization is important but challenging because the designer must choose parameters, such as auction reserve prices, to maximize a design objective while the optimal choice depends on unknown strategic behavior. The high cost of collecting strategic behavioral data limits the number of available experiments, making it difficult to simultaneously reduce behavioral uncertainty and optimize mechanism parameters. We propose a framework to address this. First, we leverage large language models to generate a pool of candidate parametric behavioral models, without requiring it to contain the true behavioral model. This converts open-ended behavioral uncertainty into a finite pool of behavioral hypotheses. Given this pool, we develop two sequential experimental design solutions to optimize mechanism parameters. The first follows a two-stage approach that selects experiments solely to reduce uncertainty about the behavioral model pool before optimizing the mechanism parameter. The second takes an end-to-end approach that selects experiments to directly reduce uncertainty about the optimal mechanism parameter. To reduce computational complexity, we approximate this experiment selection criterion using a surrogate parameter space. We further provide a theoretical guarantee that the approximation preserves the ranking of experimental designs under suitable conditions. Experiments in two strategic scenarios (static first-price auctions and dynamic pay-per-bid auctions) demonstrate that our methods consistently outperform multiple classes of baselines in mechanism parameter optimization.

1 INTRODUCTION

Mechanism parameter optimization is important in strategic environments (e.g., auctions), where a designer aims to maximize a design objective (e.g., revenue) by choosing appropriate mechanism parameters (Riley and Samuelson, 1981; Mohri and Medina, 2016). Optimizing these parameters requires understanding how participants respond strategically. For example, in a first-price auction, the item can be sold only when the highest bid exceeds the reserve price. When a high reserve price increases revenue by raising the payment floor, it may also leave the item unsold by excluding bidders.

Conventional approaches optimize mechanism parameters under a predefined agent behavioral model (e.g., an equilibrium model) (Crawford and Iriberri, 2007; Kahneman and Tversky, 2013). Nevertheless, manually specifying a model that captures all factors influencing human behavior is difficult (Hartford et al., 2016), leading to suboptimal mechanism parameter choices. Recent data-driven approaches use machine learning models either to learn strategic behavioral responses (Wang and Yu, 2024; Qian and Yu, 2026) or to directly optimize the mechanism from behavioral observations (Mohri and Medina, 2016; Dütting et al., 2019). However, these methods require large amounts of strategic behavioral data. Such data are costly to collect because each observation requires running an actual experiment with strategic participants under a chosen experimental design.

We study mechanism parameter optimization when the true behavioral model is unknown and observations of agent behavior are limited. The core challenge is that behavioral uncertainty directly induces uncertainty about the optimal mechanism parameter, yet limited observations make it infea-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

sible to reduce both simultaneously. Our key idea is to convert open-ended behavioral uncertainty into uncertainty over a finite pool of plausible behavioral hypotheses, and then reduce it through sequential experimental design to identify the optimal mechanism parameter. First, we leverage large language models (LLMs) to generate a pool of plausible parametric behavioral models, without requiring recovery of the true behavioral model. Second, given this model pool, we develop two sequential design strategies to optimize mechanism parameters. Since reducing behavioral uncertainty naturally improves mechanism parameter optimization, we first develop a two-stage approach that selects designs to identify the model that best accounts for agents' responses, and then optimizes mechanism parameters. However, when the model pool may not contain the true model and behavioral data are limited, not all reductions in behavioral uncertainty improve optimization. We therefore develop an end-to-end approach that directly reduces uncertainty about the optimal mechanism parameter. We show that this end-to-end objective, despite its apparent intractability, admits an approximation via a surrogate parameter space that provably preserves the ranking of experimental designs under suitable conditions.

Our contributions are threefold. **(i) A Mechanism Parameter Optimization Framework**: When the true behavioral model is unknown, our framework uses an LLM-generated behavioral model pool to build the uncertainty space for behavioral models and optimize the mechanism parameter through sequential experimental design. **(ii) Two Sequential Experimental Design Strategies**: We develop two sequential experimental design strategies. For the strategy directly targeting the optimal mechanism parameter, we derive a tractable approximation with a theoretical guarantee on ranking preservation under suitable conditions. **(iii) Evaluation in Strategic Scenarios**: We evaluate our framework in two strategic scenarios against multiple baselines by comparing the design objective. The results show that our framework consistently improves mechanism parameter optimization.

2 Related Work

Mechanism Parameter Optimization under Model Uncertainty Classical mechanism parameter optimization typically starts from a specified model of agent behavior, such as an equilibrium model or a parametric behavioral model, and then chooses the parameter that maximizes the designer's objective under that assumption (Myerson, 1981; Crawford and Iriberri, 2007; Platt et al., 2013). A separate line of work on automated mechanism design does not assume a fully known agent prior, and instead learns mechanisms directly from samples (Mohri and Medina, 2016; Dütting et al., 2019). However, collecting such strategic behavior data is costly, making these methods less suitable when data are limited. Our work differs from both lines of work. We do not assume any predefined model of agent behavior. Instead, we leverage LLMs to construct a candidate model pool and reduce model uncertainty using limited behavioral data.

Sequential Experimental Design Sequential experimental design adaptively selects each experiment based on previous experiment outcomes (Myung et al., 2013). Sequential Bayesian experimental design formalizes this as optimization of expected information gain (Lindley, 1956; Chaloner and Verdinelli, 1995; Foster et al., 2019; Kleinegesse and Gutmann, 2020; Zheng et al., 2020). Our work is most closely related to a line of work on goal-oriented experimental design (Filstroff et al., 2024; Huang et al., 2024; Zhong et al., 2026; Zhang et al., 2026). They select experiments with awareness of the downstream target rather than solely identifying all unknown parameters. However, existing methods rely on a known response model, such as a physical model, which characterizes how different experimental designs affect observed outcomes (Attia et al., 2018; Zhong et al., 2026), or require sufficient offline design-outcome trajectories (Huang et al., 2024; Zhang et al., 2026). In contrast, our setting involves no known model of agents' strategic behavior and limited behavioral data. We leverage LLMs to generate a pool of plausible candidate behavioral models. Our framework can optimize mechanism parameters without requiring the pool to contain the true behavioral model.

LLMs for Strategic Behavior Modeling Recent studies have examined the reasoning abilities of LLMs, showing that they can perform multi-step symbolic and logical reasoning under suitable prompting (Wei et al., 2022; Kojima et al., 2022; Liu et al., 2026), and some work has further explored using LLMs to generate models to describe human decisions (Rmus et al., 2025; Xie et al., 2026). One of the most closely related works is Xie et al. (2026), which uses LLMs to build behavioral models from human strategic data to find the model that best explains strategic behavior. We

use LLMs to generate plausible candidate behavioral models without relying on observed behavioral data, and our framework does not require recovering the true agent model.

3 Problem Formulation

We consider a sequential experimental design problem for optimizing mechanism parameters (e.g., reserve prices in first-price auctions) under behavioral uncertainty. In particular, we do not assume that strategic agents are fully rational, but instead only assume that their behavior is governed by an unknown behavioral model. Since the optimal mechanism parameter depends critically on agent behavior, determining it requires learning about the underlying behavioral model. Our goal is therefore to design a sequence of experiments that tests different mechanism parameters, learns about the agents' behavioral model based on their responses, and thereby determines the optimal mechanism parameter.

Agent Behavior Model Each behavioral model captures a distinct behavioral assumption (e.g., risk aversion or bounded rationality) and is parameterized by a model-specific behavioral parameter. We denote the true behavioral model and the corresponding true behavioral parameter by ϕ_{true} and $\boldsymbol{\theta}_{\text{true}}$, respectively. We consider a potentially large candidate pool of behavioral models, denoted by $\{\phi^{(k)}\}_{k=1}^{K}$, where each $\phi^{(k)}$ is associated with a model-specific parameter $\boldsymbol{\theta}^{(k)}$ ranging over a parameter space $\Theta^{(k)}$. Crucially, we allow for model misspecification: the candidate pool need not contain the true behavioral model. Initially, we only know a prior distribution over the candidate model and the associated parameter. We use $p\left(\phi = \phi^{(k)}\right)$ to denote the prior mass assigned to model $\phi^{(k)}$, and $p\left(\boldsymbol{\theta}^{(k)} \mid \phi = \phi^{(k)}\right)$ to denote the conditional prior over the behavioral parameters. These together induce the joint prior $p(\phi^{(k)}, \boldsymbol{\theta}^{(k)}) = p(\phi^{(k)})\, p(\boldsymbol{\theta}^{(k)} \mid \phi^{(k)})$. For brevity, we let $(\phi, \boldsymbol{\theta})$ denote the model-parameter pair. When $\phi = \phi^{(k)}$, the associated parameter is $\boldsymbol{\theta} = \boldsymbol{\theta}^{(k)} \in \Theta^{(k)}$.

Sequential Experiments Since the true behavioral model and its parameters are unknown, reducing uncertainty about them will help in reducing uncertainty about the optimal mechanism parameter. The uncertainty can be reduced if we have behavioral evidence from observations. This is achieved by conducting experiments with different mechanism parameters and observing agents' responses. However, it is costly to deploy experiments in strategic environments (e.g., running auctions with real participants), inducing a limited experimental budget. Here, we formalize the budget constraint as a limit of T sequential experiments. In each round $t = 1, \ldots, T$, the designer conducts an experiment $\boldsymbol{e}_t \in \mathcal{E}$ and observes an outcome $\boldsymbol{y}_t^{\text{obs}}$, which provides evidence about agent behavior. For instance, in a first-price auction, $\boldsymbol{e}_t$ may correspond to the reserve price, while $\boldsymbol{y}_t^{\text{obs}}$ denotes the bids submitted by the agents. Formally, the observed outcome is drawn from the distribution induced by the true behavioral model:

$$\boldsymbol{y}_t^{\text{obs}} \sim p(\boldsymbol{y} \mid \boldsymbol{e}_t, \phi_{\text{true}}, \boldsymbol{\theta}_{\text{true}}), \tag{1}$$

which characterizes how agents governed by ϕ_{true} with parameters $\boldsymbol{\theta}_{\text{true}}$ respond to experiment $\boldsymbol{e}_t$. Let $\mathcal{D}_t = \{(\boldsymbol{e}_\tau, \boldsymbol{y}_\tau^{\text{obs}})\}_{\tau=1}^{t}$ denote the dataset collected up to round t, with $\mathcal{D}_0 = \varnothing$. Given $\mathcal{D}_t$, we maintain a *posterior belief* over candidate model-parameter pairs $p(\phi, \boldsymbol{\theta} \mid \mathcal{D}_t)$. As new experimental outcomes are collected, this posterior is updated sequentially via Bayes' rule at each round.

Experimental Design Problem In each round t, the experiment $\boldsymbol{e}_t$ is chosen adaptively according to a design strategy π, such that $\boldsymbol{e}_t = \pi(\mathcal{D}_{t-1})$. The strategy π aims to reduce behavioral uncertainty while remaining informative for identifying the optimal mechanism parameter. Specifically, let $f(\boldsymbol{z}, \phi, \boldsymbol{\theta})$ denote the design objective (e.g., expected revenue) evaluated under mechanism parameter $\boldsymbol{z} \in \mathcal{Z}$ when agents behave according to model ϕ with parameter $\boldsymbol{\theta}$. After collecting the dataset $\mathcal{D}_T$ over T rounds using the strategy π, we solve

$$\hat{\boldsymbol{z}}_T = \arg\max_{\boldsymbol{z} \in \mathcal{Z}} \; \mathbb{E}_{p(\phi, \boldsymbol{\theta} \mid \mathcal{D}_T)} \left[f(\boldsymbol{z}, \phi, \boldsymbol{\theta}) \right]. \tag{2}$$

Here, $\hat{\boldsymbol{z}}_T$ is the *belief-conditioned* optimal mechanism parameter, given the current belief of agent behavior encoded in $\mathcal{D}_T$. The quality of $\hat{\boldsymbol{z}}_T$ is measured by the gap function:

$$G(\hat{\boldsymbol{z}}_T) = \max_{\boldsymbol{z} \in \mathcal{Z}} f(\boldsymbol{z}, \phi_{\text{true}}, \boldsymbol{\theta}_{\text{true}}) - f(\hat{\boldsymbol{z}}_T, \phi_{\text{true}}, \boldsymbol{\theta}_{\text{true}}), \tag{3}$$

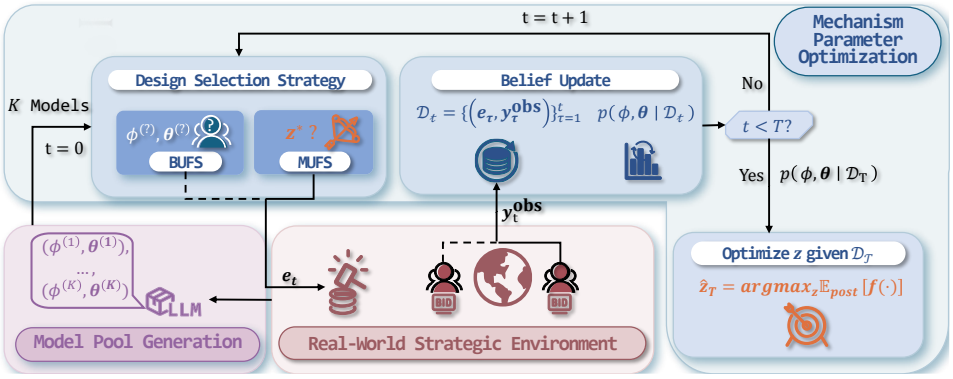


Figure 1: The pipeline of our framework.

where the first term is the optimal objective value under complete knowledge of the true behavioral model, and the second term is the objective value achieved by the selected parameter $\hat{\boldsymbol{z}}_T$. Our research question is therefore how to design a strategy π to collect an informative dataset $\mathcal{D}_T$ that minimizes this gap.

4 Our Framework

We develop a framework for the sequential experimental design problem formulated in Section 3. The framework consists of two components. First, we use an LLM to generate a candidate pool of behavioral models. Second, given this pool, we consider two solution approaches: (i) a two-stage solution instantiated by the *Behavioral Uncertainty-First Strategy* (**BUFS**), which first reduces uncertainty about the behavioral model and its associated parameters and then optimizes the mechanism parameter based on the resulting posterior belief; and (ii) an end-to-end solution instantiated by the *Mechanism Uncertainty-First Strategy* (**MUFS**), which directly reduces uncertainty about the optimal mechanism parameter. The overall pipeline is illustrated in Figure 1.

4.1 LLM-Generated Candidate Behavioral Model Pool

Motivation for Using LLMs As established in Section 3, we do not assume that strategic agents are fully rational. Under such behavioral uncertainty, mechanism parameter optimization requires a rich candidate pool $\{\phi^{(k)}\}_{k=1}^{K}$ that can represent a broad range of plausible agent behaviors. Constructing such a pool manually relies heavily on domain expertise and often results in a narrow set of classical models. To overcome this limitation, we use an LLM to construct the candidate pool. Recovering the true behavioral model ϕ_{true} with an LLM is generally difficult, and our framework does not rely on this being possible. Instead, the key requirement is that the LLM-generated pool be sufficiently expressive so that it can approximate the true behavior-generating process in (1).

This requirement is well aligned with the strengths of modern LLMs. Pretrained on large corpora spanning economics, game theory, and psychology, they are well suited to retrieving and recombining behavioral knowledge relevant to a given mechanism. As a result, they can span a broad behavioral hypothesis space and produce informative approximations even when the true behavior does not coincide with any canonical textbook model (Xie et al., 2026).

Behavioral Model Pool Construction We prompt the LLM with a textual description of the mechanism rules and payoff structure to generate the behavioral model pool. Each proposed model $\phi^{(k)}$ is returned together with a model-specific parameter $\boldsymbol{\theta}^{(k)} \in \Theta^{(k)}$, thereby defining a functional form of agent behavior, as well as a human-readable interpretation that makes the underlying assumptions readily inspectable. Each candidate model $\phi^{(k)}$ with parameter $\boldsymbol{\theta}^{(k)}$ specifies the distribution of agents' responses $\boldsymbol{y}$ given an experiment $\boldsymbol{e}$.

4.2 A TWO-STAGE SOLUTION FOR MECHANISM OPTIMIZATION VIA **BUFS**

Given the candidate pool $\{\phi^{(k)}\}_{k=1}^{K}$ generated by the LLM, we introduce a two-stage solution for mechanism optimization instantiated by **BUFS**. With T experiments, **BUFS** first sequentially selects experiments to reduce uncertainty about the candidate model-parameter pair, thereby updating the posterior belief toward behavioral hypotheses that best explain the observed agent responses. It then computes the belief-conditioned optimal mechanism parameter $\hat{\boldsymbol{z}}_T$ according to (2).

Stage 1: Behavioral Uncertainty Reduction via BUFS The first stage aims to reduce uncertainty about the model-parameter pair that best approximates the agents' responses. Therefore, **BUFS** is based on the standard sequential Bayesian optimal experimental design (SBOED) (Zheng et al., 2020). Specifically, at each round t, **BUFS** selects the experiment that maximizes the Expected Information Gain (EIG) about $(\phi, \boldsymbol{\theta})$ given the historical dataset $\mathcal{D}_{t-1}$. Formally, let $\boldsymbol{y}_t$ denote the outcome variable at round t, let $\boldsymbol{e}$ denote the experimental design variable at round t, and let $H(\cdot)$ denote the entropy. Then,

$$\begin{aligned}\boldsymbol{e}_t &= \arg\max_{\boldsymbol{e}\in\mathcal{E}} \ \mathbb{E}_{p(\boldsymbol{y}_t|\boldsymbol{e},\mathcal{D}_{t-1})}\Big(H(p(\phi,\boldsymbol{\theta}\mid\mathcal{D}_{t-1})) - H(p(\phi,\boldsymbol{\theta}\mid\boldsymbol{e},\boldsymbol{y}_t,\mathcal{D}_{t-1}))\Big)\\ &= \arg\max_{\boldsymbol{e}\in\mathcal{E}} \ \mathbb{E}_{p(\phi,\boldsymbol{\theta},\boldsymbol{y}_t|\boldsymbol{e},\mathcal{D}_{t-1})}\left(\log\frac{p(\boldsymbol{y}_t\mid\boldsymbol{e},\phi,\boldsymbol{\theta},\mathcal{D}_{t-1})}{p(\boldsymbol{y}_t\mid\boldsymbol{e},\mathcal{D}_{t-1})}\right).\end{aligned} \tag{4}$$

Equation (4) shows that **BUFS** favors experiments that maximally reduce the expected uncertainty over $(\phi, \boldsymbol{\theta})$. Since $p(\boldsymbol{y}_t \mid \boldsymbol{e}, \mathcal{D}_{t-1})$ in (4) usually lacks a closed form, the objective in (4) is commonly approximated with a nested Monte Carlo estimator (Foster et al., 2019; Myung et al., 2013).

Stage 2: Mechanism Optimization under the Updated Posterior After T rounds of experiments, the posterior $p(\phi, \boldsymbol{\theta} \mid \mathcal{D}_T)$ is expected to become increasingly concentrated on the models and parameter regions that best explain agent behavior governed by $(\phi_{\text{true}}, \boldsymbol{\theta}_{\text{true}})$. Based on this posterior belief, we solve for the belief-conditioned optimal mechanism parameter $\hat{\boldsymbol{z}}_T$ according to (2). In this way, the two-stage solution is intended to reduce the gap $G(\hat{\boldsymbol{z}}_T)$ in (3).

4.3 AN END-TO-END SOLUTION FOR MECHANISM OPTIMIZATION VIA **MUFS**

While **BUFS** reduces uncertainty about the model-parameter pair $(\phi, \boldsymbol{\theta})$, not all behavioral uncertainty affects the mechanism optimization objective. For instance, two candidate models may differ substantially in their behavioral assumptions yet induce nearly identical optimal mechanism parameters. In such cases, **BUFS** may waste experimental budget on experiments that carry no value for the final objective. Therefore, we introduce an end-to-end solution for mechanism optimization via **MUFS** that addresses this by directly reducing uncertainty about the optimal mechanism parameter.

Mechanism Uncertainty-First Strategy Let $g(\phi, \boldsymbol{\theta}) = \arg\max_{\boldsymbol{z}\in\mathcal{Z}} f(\boldsymbol{z}, \phi, \boldsymbol{\theta})$ denote the mapping from a behavioral model-parameter pair to its optimal mechanism parameters. We define $\boldsymbol{z}^* = g(\phi, \boldsymbol{\theta})$ as the corresponding induced optimal mechanism parameter variable. In each round t, **MUFS** selects the experiment that maximizes the expected information gain about the optimal mechanism parameters $\boldsymbol{z}^*$ given the historical dataset $\mathcal{D}_{t-1}$. Formally, $\boldsymbol{e}_t$ is selected according to

$$\begin{aligned}\boldsymbol{e}_t &= \arg\max_{\boldsymbol{e}\in\mathcal{E}} \mathbb{E}_{p(\boldsymbol{y}_t|\boldsymbol{e},\mathcal{D}_{t-1})}(H(p(\boldsymbol{z}^*\mid\mathcal{D}_{t-1})) - H(p(\boldsymbol{z}^*\mid\boldsymbol{e},\boldsymbol{y}_t,\mathcal{D}_{t-1})))\\ &= \arg\max_{\boldsymbol{e}\in\mathcal{E}} \mathbb{E}_{p(\boldsymbol{z}^*,\boldsymbol{y}_t|\boldsymbol{e},\mathcal{D}_{t-1})}\left(\log\frac{p(\boldsymbol{z}^*\mid\boldsymbol{e},\boldsymbol{y}_t,\mathcal{D}_{t-1})}{p(\boldsymbol{z}^*\mid\mathcal{D}_{t-1})}\right).\end{aligned} \tag{5}$$

Equation (5) shows that **MUFS** evaluates experiments according to how much they are expected to reduce uncertainty about $\boldsymbol{z}^*$. The uncertainty about $\boldsymbol{z}^*$ arises from the uncertainty about $(\phi, \boldsymbol{\theta})$, since the optimal parameter vector is computed directly from the behavior model and its associated parameters. Equivalently, the objectives of **BUFS** and **MUFS** can be written as the conditional mutual information $I(\phi, \boldsymbol{\theta}; \boldsymbol{y}_t \mid \boldsymbol{e}, \mathcal{D}_{t-1})$ and $I(\boldsymbol{z}^*; \boldsymbol{y}_t \mid \boldsymbol{e}, \mathcal{D}_{t-1})$, respectively. Given that $\boldsymbol{z}^*$ is determined by $(\phi, \boldsymbol{\theta})$, the distinction between **BUFS** and **MUFS** can be further characterized by

$$I(\phi, \boldsymbol{\theta}; \boldsymbol{y}_t \mid \boldsymbol{e}, \mathcal{D}_{t-1}) = I(z^*; \boldsymbol{y}_t \mid \boldsymbol{e}, \mathcal{D}_{t-1}) + I(\phi, \boldsymbol{\theta}; \boldsymbol{y}_t \mid \boldsymbol{z}^*, \boldsymbol{e}, \mathcal{D}_{t-1}). \tag{6}$$

Thus, unlike **BUFS**, **MUFS** selects experiments that most sharpen posterior belief about the optimal mechanism parameter $\boldsymbol{z}^*$, without additionally valuing information about $(\phi, \boldsymbol{\theta})$ beyond what is needed to identify $\boldsymbol{z}^*$. This aligns the selection of experiments with the final mechanism optimization objective. See Appendix A.1 for further discussion.

Note that $H(p(\boldsymbol{z}^* \mid \mathcal{D}_{t-1}))$ does not depend on the experiment design $\boldsymbol{e}$. We rewrite (5) by expanding $p(\boldsymbol{z}^*, \boldsymbol{y}_t \mid \boldsymbol{e}, \mathcal{D}_{t-1})$ by marginalizing over the candidate model-parameter pairs:

$$\boldsymbol{e}_t = \arg\max_{\boldsymbol{e}\in\mathcal{E}} \int_{\boldsymbol{y}_t} \int_{\boldsymbol{z}^*} \sum_{k=1}^{K} \int_{\boldsymbol{\theta}^{(k)}} \log(p(\boldsymbol{z}^* \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1}))\, p(\boldsymbol{z}^*, \boldsymbol{y}_t, \phi^{(k)}, \boldsymbol{\theta}^{(k)} \mid \boldsymbol{e}, \mathcal{D}_{t-1}) d\boldsymbol{\theta}^{(k)} d\boldsymbol{z}^* d\boldsymbol{y}_t, \quad (7)$$

where the joint term can be factorized as

$$p(\boldsymbol{z}^*, \boldsymbol{y}_t, \phi^{(k)}, \boldsymbol{\theta}^{(k)} \mid \boldsymbol{e}, \mathcal{D}_{t-1}) = p(\boldsymbol{y}_t \mid \boldsymbol{e}, \mathcal{D}_{t-1}) p(\phi^{(k)}, \boldsymbol{\theta}^{(k)} \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1}) p(\boldsymbol{z}^* \mid \phi^{(k)}, \boldsymbol{\theta}^{(k)}). \quad (8)$$

Compared with the experiment selection criterion in (4) in **BUFS**, the above selection criterion is substantially harder to evaluate. The key challenge lies in computing $p(\boldsymbol{z}^* | \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1})$ and $p\left(\boldsymbol{z}^* | \phi^{(k)}, \boldsymbol{\theta}^{(k)}\right)$. Recall that $\boldsymbol{z}^*$ denotes the solution to a design objective maximization problem. Consequently, evaluating the above distributions requires solving this optimization problem for each candidate model-parameter pair. Since the model-parameter space is high-dimensional, this results in a large number of optimization solves, making the overall computation particularly challenging.

Selection Criterion Approximation via Surrogates To make **MUFS** tractable, we introduce a surrogate space $\tilde{\mathcal{Z}}$ for the original space $\mathcal{Z}$. Specifically, we construct $\tilde{\mathcal{Z}}$ by quantizing the elements of $\mathcal{Z}$ on a uniform grid. We define $\tilde{\boldsymbol{z}}^*$ as the optimal solution in $\tilde{\mathcal{Z}}$ that maximizes the design objective. Given the behavior model and its associated parameters, computing $\tilde{\boldsymbol{z}}^*$ is much simpler than computing $\boldsymbol{z}^*$, since it only requires direct comparison of the objective values over the elements of $\tilde{\mathcal{Z}}$. Therefore, rather than selecting the design $\boldsymbol{e}$ to reduce uncertainty about $\boldsymbol{z}^*$ as in (7), we select $\boldsymbol{e}$ to reduce uncertainty about $\tilde{\boldsymbol{z}}^*$. This can significantly reduce the computational cost. After replacing $\boldsymbol{z}^*$ with $\tilde{\boldsymbol{z}}^*$, we approximate the resulting criterion using a nested Monte Carlo estimator:

$$\boldsymbol{e}_t = \arg\max_{\boldsymbol{e}\in\mathcal{E}} \frac{1}{N} \sum_{n=1}^{N} \left(\frac{1}{M} \sum_{m=1}^{M} \log \left(\frac{1}{J} \sum_{j=1}^{J} \mathbb{I}\big(\tilde{\boldsymbol{z}}^*_{n,m} = \tilde{\boldsymbol{z}}^*_{n,m,j}\big) \right) \right). \quad (9)$$

The estimation procedure consists of three sampling layers. The outer layer samples $(\phi_n, \boldsymbol{\theta}_n) \sim p(\phi, \boldsymbol{\theta} \mid \mathcal{D}_{t-1})$ and prospective outcomes $\boldsymbol{y}_n \sim p(\boldsymbol{y} \mid \phi_n, \boldsymbol{\theta}_n, \boldsymbol{e}, \mathcal{D}_{t-1})$. Given $\boldsymbol{y}_n$, the middle layer samples $(\phi_{n,m}, \boldsymbol{\theta}_{n,m}) \sim p(\phi, \boldsymbol{\theta} \mid \boldsymbol{e}, \boldsymbol{y}_n, \mathcal{D}_{t-1})$ and computes $\tilde{\boldsymbol{z}}^*_{n,m} = \arg\max_{\tilde{\boldsymbol{z}}\in\tilde{\mathcal{Z}}} f(\tilde{\boldsymbol{z}}, \phi_{n,m}, \boldsymbol{\theta}_{n,m})$. The inner layer samples $(\phi_{n,m,j}, \boldsymbol{\theta}_{n,m,j}) \sim p(\phi, \boldsymbol{\theta} \mid \boldsymbol{e}, \boldsymbol{y}_n, \mathcal{D}_{t-1})$ and computes $\tilde{\boldsymbol{z}}^*_{n,m,j} = \arg\max_{\tilde{\boldsymbol{z}}\in\tilde{\mathcal{Z}}} f(\tilde{\boldsymbol{z}}, \phi_{n,m,j}, \boldsymbol{\theta}_{n,m,j})$. Then, it approximates $\log\left(p(\tilde{\boldsymbol{z}}^*_{n,m} \mid \boldsymbol{e}, \boldsymbol{y}_n, \mathcal{D}_{t-1})\right)$ as $\log\left(\frac{1}{J}\sum_{j=1}^{J} \mathbb{I}\big(\tilde{\boldsymbol{z}}^*_{n,m} = \tilde{\boldsymbol{z}}^*_{n,m,j}\big)\right)$.

We show that, under suitable conditions, a sufficiently fine uniform discretization of $\mathcal{Z}$ can preserve the ranking of designs and serve as a reliable surrogate for comparing experiment designs under **MUFS**. Specifically, we let $\mathcal{Z} \subset \mathbb{R}^d$ be a compact rectangular space. We partition $\mathcal{Z}$ into uniform grid cells $\{\mathcal{Z}_i^h\}_{i=1}^I$, each with side length h and volume h^d, and let $\tilde{\boldsymbol{z}}_i$ be the midpoint of $\mathcal{Z}_i^h$. Define $\tilde{\mathcal{Z}} = Q_h(\mathcal{Z}) = \{\tilde{\boldsymbol{z}}_1, \ldots, \tilde{\boldsymbol{z}}_I\}$ and let $\mathcal{E}$ be finite. We introduce the following proposition (see Appendix A.2 for the proof).

Proposition 1. *Given $(\phi, \boldsymbol{\theta})$, suppose that $\boldsymbol{z}^*$ is unique. Assume that if $\boldsymbol{z}^* \in \mathcal{Z}_i^h$, then $\tilde{\boldsymbol{z}}^* = \tilde{\boldsymbol{z}}_i$, indicating that $\tilde{\boldsymbol{z}}^*$ is the midpoint of the cell that contains $\boldsymbol{z}^*$. Suppose there exist constants $m > 0$ and $G < \infty$ such that $p(\boldsymbol{z}^* \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1})$ is lower bounded by m and is G-Lipschitz on $\mathcal{Z}$ for all $\boldsymbol{e}, \boldsymbol{y}_t, \boldsymbol{z}^*$. Define $\mathcal{S}(\boldsymbol{e}) := \mathbb{E}_{p(\boldsymbol{y}_t|\boldsymbol{e},\mathcal{D}_{t-1})}[H(p(\boldsymbol{z}^* \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1}))]$, $\tilde{\mathcal{S}}(\boldsymbol{e}) := \mathbb{E}_{p(\boldsymbol{y}_t|\boldsymbol{e},\mathcal{D}_{t-1})}[H(p(\tilde{\boldsymbol{z}}^* \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1})) + d \log h]$, and*

$$\Delta := \min_{\substack{\boldsymbol{e}_1, \boldsymbol{e}_2 \in \mathcal{E} \\ \mathcal{S}(\boldsymbol{e}_1) \neq \mathcal{S}(\boldsymbol{e}_2)}} |\mathcal{S}(\boldsymbol{e}_1) - \mathcal{S}(\boldsymbol{e}_2)|. \quad (10)$$

Let $A_h := Ih^d$ denote the total volume of the grid cells. If $\Delta > 0$ and $h < \sqrt{24m\Delta/(dG^2A_h)}$, then the strict pairwise rankings of experiments induced by $\tilde{\boldsymbol{z}}^*$ are preserved under $\boldsymbol{z}^*$.

5 EXPERIMENTS

We evaluate our framework on mechanism parameter optimization against several baselines, using the metric defined in (3). We consider two complementary strategic games: a static first-price auction (Rosenkranz and Schmitz, 2007) and a dynamic pay-per-bid auction (Platt et al., 2013; Wang and Yu, 2024).

Use of LLMs in the Experiments We use LLMs in the experiments for two different purposes. First, as described in Section 4.1, we prompt LLMs to generate the candidate model pool. Second, we use LLM-based agents as strategic decision-makers (e.g., bidders in auctions) and collect the resulting game outcomes. Using LLM-based agents does not assume that they faithfully replicate human behavior. Rather, it is motivated by recent studies showing that LLM-based agents exhibit nontrivial strategic decision-making behavior in game-theoretic and multi-agent scenarios (Duan et al., 2024; Herr et al., 2024; Mozikov et al., 2024; Mao et al., 2025). Therefore, we use them as controllable agents that can respond systematically to changes in mechanism parameters and adjust their decisions accordingly, displaying predictable strategic patterns. This makes them a practical testbed for studying mechanism parameter optimization under strategic behavioral responses.

In implementation, each LLM-based agent receives the decision context (e.g., the mechanism description, the payoff structure) without being instructed to follow any particular strategy. We test agents based on different LLMs: an open-source model, DeepSeek-V3.1 (Liu et al., 2024), and a closed-source model, GPT-5.1 (OpenAI, 2025). Details are provided in Appendix C.5.

Comparison Methods We compare our **BUFS** and **MUFS** with four categories of baselines.

- **Sequential Design Methods**: They share the same sequential experimental design principle as ours but use different design selection strategies. **Seq-Random** selects designs randomly. **Seq-HC** selects designs via hill-climbing that greedily moves to a neighboring design with higher revenue (Selman and Gomes, 2006). **Seq-TS** selects designs via Thompson Sampling that greedily picks the best given current posterior belief to the model-parameter (Chapelle and Li, 2011).
- **Bandit-Based Methods**: They learn how mechanism parameters affect the design objective from sequential experimental observations and select subsequent parameters by balancing exploration and exploitation, without relying on behavioral assumptions. They include **LinTS** (Agrawal and Goyal, 2013) and **Epoch Greedy** (Langford and Zhang, 2007).
- **Search-Based Methods**: They treat mechanism optimization as a black-box search, selecting designs at each round using outcomes from past designs. They do not maintain a posterior over behavioral models and return the best-observed mechanism parameter. They include **Heuristic Search** (Russell and Norvig, 2020), **Coarse Search** (Dudek, 2010), and **Random Sampling** (Bergstra and Bengio, 2012).
- **Equilibrium-Based Method**: It assumes that agents play an equilibrium strategy and directly optimizes the mechanism parameters based on this known behavior model, without leveraging information from sequential experiments (Myerson, 1981; Platt et al., 2013).

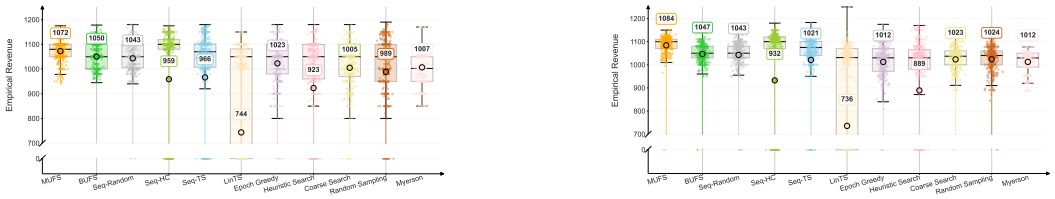
All methods use the same experimental budget, and reported results are averaged over 10 independent runs with different random seeds. In each run, the optimized mechanism parameter is evaluated over 30 auctions, yielding 300 auctions per method.

5.1 RESERVE PRICE OPTIMIZATION IN THE FIRST-PRICE AUCTION

Experimental Setup We consider a first-price auction with I bidders and a public reserve price z. Each bidder has a private value independently drawn from a uniform distribution, where the distribution and the number of bidders are common knowledge. Each bidder submits a sealed bid for the single item. The bidder with the highest bid wins and pays their bid if the bid exceeds the reserve price z; otherwise, the item remains unsold. By screening out bids below z, the reserve price determines allocation and payment, making it a natural mechanism parameter to optimize for maximizing the seller's revenue. In the first-price auction, what we can observe is the bid vector

Table 1: Empirical revenue comparison over 300 first-price auctions under different LLM-based agents. The best values are in **bold**, and the second-best values are underlined.

Method		MUFS	BUFS	Seq-Random	Seq-HC	Seq-TS	LinTS	Epoch Greedy	Heuristic Search	Coarse Search	Random Sampling	Myerson
DeepSeek-V3.1	Mean	**1072.5**	1050.4	1043.3	958.8	966.3	743.6	1022.8	922.9	1005.2	988.7	1006.9
	Std.	**44.2**	48.6	98.9	377.0	326.5	487.8	127.4	331.5	164.6	228.2	70.2
GPT-5.1	Mean	**1084.2**	1046.6	1042.6	932.5	1020.5	735.7	1012.4	889.0	1023.4	1024.0	1012.4
	Std.	70.2	70.4	112.1	402.6	233.6	499.7	105.1	371.8	96.8	78.9	**45.8**



(a) DeepSeek-V3.1-based agents (b) GPT-5.1-based agents

Figure 2: Distribution of empirical revenue over 300 first-price auctions under different agents.

$\boldsymbol{y} = (y_1, \ldots, y_I)$ of I bidders. Given $(\phi, \boldsymbol{\theta})$, y_i is bidder i's bid drawn according to (1). Thus, the design objective is the seller's expected revenue

$$f(z, \phi, \boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{y} \sim p(\boldsymbol{y}|z,\phi,\boldsymbol{\theta})} \left[\max_{i=1,\ldots,I} y_i \cdot \mathbb{I}(y_i \geq z) \right]. \tag{11}$$

Comparison Results Table 1 reports the empirical revenue obtained using DeepSeek-V3.1 and GPT-5.1. For each LLM, we conduct experiments with 10 agents. In the first-price auction, the **Equilibrium-Based Method** corresponds to **Myerson** baseline, which derives the optimal reserve price analytically under the equilibrium model via Myerson's optimal auction theory (Myerson, 1981). **MUFS** and **BUFS** rank first and second in mean revenue across different LLMs, showing their effectiveness in improving the design objective. Meanwhile, **Myerson**, **MUFS**, and **BUFS** all achieve low standard deviations, indicating stable performance across auctions.

Figure 2 shows the distribution of empirical revenue across methods. Each small point denotes the revenue from a single auction under the reserve price optimized by the corresponding method. The larger outlined point denotes the mean revenue. A zero-revenue outcome indicates a severe failure where the optimized reserve price is too high for any bid to exceed, so the item remains unsold. Frequent occurrences of such outcomes signal that a method is unreliable in optimizing the reserve price. In Figure 2a, **MUFS**, **BUFS**, and **Myerson** produce no zero-revenue outcomes, while all other baselines do. In Figure 2b, most baselines still exhibit multiple such outcomes, while **MUFS**, **BUFS**, and **Random Sampling** each produce only one such case out of 300 auctions. Although **Myerson** avoids this failure, its mean revenue is lower than ours. Overall, these results show that our methods outperform the baselines in reserve price optimization for first-price auctions.

5.2 Bid Fee and Bid Increment Optimization in Pay-per-Bid Auctions

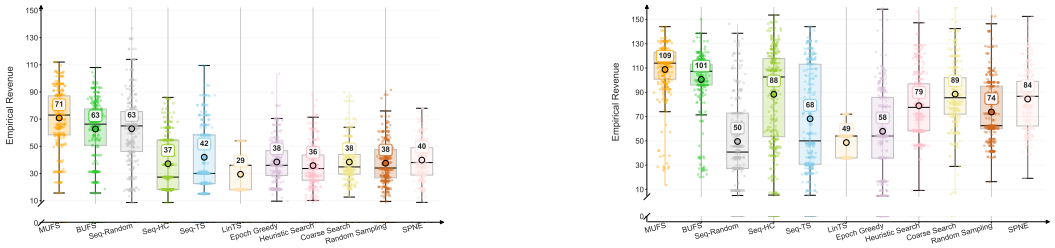
Experimental Setup We next consider a dynamic pay-per-bid auction. A bid in this auction is valid only after the bidder pays a bid fee of b. The item price starts at 0. A valid bid increases the item price by bid increment d, makes the bidder the current leader, and starts the next round. The auction ends when no one bids. The current leader wins the item and pays the final item price. If the auction lasts for n rounds, the final item price is nd. See Appendix C.1 for more introduction of the auction. Here, the mechanism parameter vector consists of the bid fee and the bid increment, denoted by $\boldsymbol{z} = (b, d)^\top$, and what we can observe is the auction duration. Given $(\phi, \boldsymbol{\theta})$, the auction duration y is drawn according to (1). Thus, the design objective is the expected revenue

$$f(\boldsymbol{z}, \phi, \boldsymbol{\theta}) = \mathbb{E}_{y \sim p(y|\boldsymbol{z},\phi,\boldsymbol{\theta})} \left[y \mathbf{1}^\top \boldsymbol{z} \right]. \tag{12}$$

Comparison Results For results in Table 2 and Figure 3, we conduct experiments with 5 agents for each LLM. In pay-per-bid auctions, the **Equilibrium-Based Method** is instantiated by the **SPNE** baseline, which analytically optimizes the mechanism parameters by assuming that bidders follow a symmetric Subgame Perfect Nash Equilibrium (Platt et al., 2013). **MUFS** achieves the highest

Table 2: Empirical revenue comparison over 300 pay-per-bid auctions under different LLM-based agents. The best values are in **bold**, and the second-best values are underlined.

Method		MUFS	BUFS	Seq-Random	Seq-HC	Seq-TS	LinTS	Epoch Greedy	Heuristic Search	Coarse Search	Random Sampling	SPNE
DeepSeek-V3.1	Mean	**70.9**	62.7	62.9	37.2	41.9	29.4	38.4	35.8	38.4	37.5	39.9
	Std.	20.9	21.5	26.4	19.7	23.2	**10.9**	15.4	14.2	14.6	15.3	14.7
GPT-5.1	Mean	**108.8**	100.6	49.5	88.3	68.2	48.6	57.9	78.9	88.5	73.8	84.5
	Std.	21.5	24.2	33.5	38.1	40.9	**10.5**	36.7	26.2	26.7	25.5	24.2



(a) DeepSeek-V3.1-based agents

(b) GPT-5.1-based agents

Figure 3: Distribution of empirical revenue over 300 pay-per-bid auctions under different agents.

mean revenue and **BUFS** also shows competitive performance. Although **Seq-Random** slightly outperforms **BUFS** in mean revenue by 0.2 under DeepSeek-V3.1, it has the highest variance among all methods. Since the pay-per-bid auction is dynamic, bidders exhibit diverse sequential strategies, making model uncertainty harder to reduce. While almost all baselines vary substantially across agents, **BUFS** and **MUFS** remain relatively stable while delivering strong average performance.

5.3 Additional Experiments

Additional experiments in Appendix B further analyze our framework from the following aspects.

Posterior Update Dynamics To examine how different design strategies reduce behavioral uncertainty, Appendix B.1 visualizes posterior beliefs over candidate models and their behavioral parameters across sequential experiments. It shows that **MUFS** and **BUFS** achieve faster posterior concentration than the sequential baselines, indicating more efficient uncertainty reduction.

Robustness to Model Pool Misspecification As noted in Section 3, our framework allows model misspecification: the candidate pool need not contain the true behavioral model of target agents. For a more direct assessment, we fit an oracle behavioral model to real human behavioral data from pay-per-bid auctions following Brünner et al. (2019). We treat the human-data-calibrated model as the ground-truth agent interacting with different experimental design methods. Appendix B.2 compares candidate pools with and without this oracle model. **MUFS** and **BUFS** maintain strong performance, even when the oracle is excluded, demonstrating robustness to model-pool misspecification.

Effect of Candidate Pool Size under Misspecification We further investigate the effect of misspecified model-pool size by exhaustively evaluating all candidate pools of sizes two and three. Appendix B.3 shows that **MUFS** and **BUFS** consistently outperform the sequential baselines under both sizes, and further improve as the pool expands. This suggests that a richer candidate pool can improve performance by providing more diverse behavioral hypotheses under misspecification.

6 Conclusion

To optimize the mechanism parameter when the true behavioral model is unknown, we proposed a framework instantiated by two sequential experimental design solutions. Experiments in two strategic scenarios show that this framework can effectively improve mechanism parameter optimization. Compared with reducing uncertainty about behavioral models, directly reducing uncertainty about optimal mechanism parameters can yield more informative data for mechanism optimization. A possible extension direction is to study how the quality of the model pool affects optimization performance. For example, increasing pool diversity may raise the likelihood of including a model close to the true one, but it may also introduce additional uncertainty that needs to be reduced.

AI USE STATEMENT

In this work, we used generative AI tools to generate synthetic data sets, assist in the writing of proofs, provide feedback on experiments, clean and reformat datasets, provide critical ingredients for proving mathematical claims and verify mathematical claims formulated by the authors. We have not used generative AI tools to help formulate the core methodological framework, propose or refine hypotheses of the framework, assist with translation, originate mathematical claims, support qualitative and thematic data analysis, help develop conceptual frameworks, implement methods, interpret results, or design research methodology.

Additionally, we used generative AI tools to suggest experimental parameters, create or edit software code, draft parts of a research paper, summarize or analyze existing literature, discover research topics or identify gaps, brainstorm, search for information, edit a research paper to improve readability, and identify relevant literature. We have reviewed all AI-assisted work. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

We have taken several steps to facilitate reproducibility. For the theoretical results, the problem formulation and algorithmic details of the proposed methods are described in Section 3 and Section 4. Appendix A.1 provides detailed derivations of the relationship between **BUFS** and **MUFS** and sufficient conditions under which their objectives coincide. C.3 shows detailed algorithmic description of **BUFS** and **MUFS**. Meanwhile, the assumptions of Proposition 1 are stated explicitly in the main text, with a complete proof provided in Appendix A.2.

For the empirical results, Section 5 describes the experimental settings, comparison methods, and evaluation results, while Appendix B reports additional analyses of posterior update dynamics, robustness to model-pool misspecification using a human-data-calibrated oracle agent, the effect of misspecified model-pool size, and evaluation beyond LLM-based agents. Further implementation details are provided in Appendix C, including the tested auction environments, hyperparameter settings, candidate behavioral models, LLM implementation settings, and prompts used in the experiments. We have also submitted an anonymized code archive as supplementary material during review, containing implementations of the proposed methods, the experimental pipeline, relevant baselines, and prompts. After acceptance, we will publicly release the code in a permanent repository.

REFERENCES

John G Riley and William F Samuelson. Optimal auctions. *The American Economic Review*, 71(3): 381–392, 1981.

Mehryar Mohri and Andrés Munoz Medina. Learning algorithms for second-price auctions with reserve. *Journal of Machine Learning Research*, 17(74):1–25, 2016.

Vincent P Crawford and Nagore Iriberri. Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions? *Econometrica*, 75 (6):1721–1770, 2007.

Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*, pages 99–127. World Scientific, 2013.

Jason S Hartford, James R Wright, and Kevin Leyton-Brown. Deep learning for predicting human strategic behavior. *Advances in neural information processing systems*, 29, 2016.

Yujia Wang and Haoran Yu. Predicting real-world penny auction durations by integrating game theory and machine learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 9926–9934, 2024.

Xingyu Qian and Haoran Yu. Inferring heterogeneous private valuations from offline market data via entropic risk-sensitive utility maximization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pages 17661–17669, 2026.

Paul Dütting, Zhe Feng, Harikrishna Narasimhan, David Parkes, and Sai Srivatsa Ravindranath. Optimal auctions through deep learning. In *Proceedings of the International Conference on Machine Learning*, pages 1706–1715. PMLR, 2019.

Roger B Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.

Brennan C Platt, Joseph Price, and Henry Tappen. The role of risk preferences in pay-to-bid auctions. *Management Science*, 59(9):2117–2134, 2013.

Jay I Myung, Daniel R Cavagnaro, and Mark A Pitt. A tutorial on adaptive design optimization. *Journal of Mathematical Psychology*, 57(3-4):53–67, 2013.

Dennis V Lindley. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 27(4):986–1005, 1956.

Kathryn Chaloner and Isabella Verdinelli. Bayesian experimental design: A review. *Statistical science*, pages 273–304, 1995.

Adam Foster, Martin Jankowiak, Elias Bingham, Paul Horsfall, Yee Whye Teh, Thomas Rainforth, and Noah Goodman. variational bayesian optimal experimental design. *Advances in neural information processing systems*, 32, 2019.

Steven Kleinegesse and Michael U Gutmann. Bayesian experimental design for implicit models by mutual information neural estimation. In *Proceedings of International Conference on Machine Learning*, pages 5316–5326. PMLR, 2020.

Sue Zheng, David Hayden, Jason Pacheco, and John W Fisher III. sequential bayesian experimental design with variable cost structure. *Advances in Neural Information Processing Systems*, 33: 4127–4137, 2020.

Louis Filstroff, Iiris Sundin, Petrus Mikkola, Aleksei Tiulpin, Juuso Kylmäoja, and Samuel Kaski. Targeted active learning for bayesian decision-making. *Transactions on Machine Learning Research*, 2024. ISSN 2835-8856.

Daolang Huang, Yujia Guo, Luigi Acerbi, and Samuel Kaski. Amortized bayesian experimental design for decision-making. *Advances in Neural Information Processing Systems*, 37:109460–109486, 2024.

Shijie Zhong, Wanggang Shen, Tommie Catanach, and Xun Huan. Goal-oriented bayesian optimal experimental design for nonlinear models using markov chain monte carlo. *SIAM/ASA Journal on Uncertainty Quantification*, 14(1):19–47, 2026.

Zheyu Zhang, Jiayuan Dong, Jie Liu, and Xun Huan. Goal-oriented sequential bayesian experimental design for causal learning, 2026.

Ahmed Attia, Alen Alexanderian, and Arvind K Saibaba. Goal-oriented optimal design of experiments for large-scale bayesian linear inverse problems. *Inverse Problems*, 34(9):095009, 2018.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35:24824–24837, 2022.

Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. Large language models are zero-shot reasoners. *Advances in neural information processing systems*, 35:22199–22213, 2022.

Wenxuan Liu, Xiyuan Zhou, Xinlei Wang, Yuheng Cheng, Lixin Ye, Randall Berry, Leandros Tassiulas, Jianwei Huang, and Junhua Zhao. An LLM agent-based framework for analytical characterization of nash equilibria. *Nexus*, 3(1), 2026.

Milena Rmus, Akshay Kumar Jagadish, Marvin Mathony, Tobias Ludwig, and Eric Schulz. Generating computational cognitive models using large language models. In D. Belgrave, C. Zhang, H. Lin, L. Montoya, R. Pascanu, P. Koniusz, M. Ghassemi, and N. Chen, editors, *Advances in Neural Information Processing Systems*, volume 38, pages 87796–87833, 2025.

Xiaohan Xie, Haoran Yu, Biying Shou, and Jianwei Huang. Automated human strategic behavior modeling via large language models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pages 17877–17885, 2026.

Stephanie Rosenkranz and Patrick W Schmitz. Reserve prices in auctions as reference points. *The Economic Journal*, 117(520):637–653, 2007.

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

Nathan Herr, Fernando Acero, Roberta Raileanu, María Pérez-Ortiz, and Zhibin Li. Are large language models strategic decision makers? a study of performance and bias in two-player non-zero-sum games. *arXiv preprint arXiv:2407.04467*, 2024.

Mikhail Mozikov, Nikita Severin, Valeria Bodishtianu, Maria Glushanina, Ivan Nasonov, Daniil Orekhov, Vladislav Pekhotin, Ivan Makovetskiy, Mikhail Baklashkin, Vasily Lavrentyev, et al. Eai: Emotional decision-making of LLMs in strategic games and ethical dilemmas. *Advances in Neural Information Processing Systems*, 37:53969–54002, 2024.

Shaoguang Mao, Yuzhe Cai, Yan Xia, Wenshan Wu, Xun Wang, Fengyi Wang, Qiang Guan, Tao Ge, and Furu Wei. Alympics: LLM agents meet game theory. In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 2845–2866, 2025.

Aixin Liu, Bei Feng, Bing Xue, Bingxuan Wang, Bochao Wu, Chengda Lu, Chenggang Zhao, Chengqi Deng, Chenyu Zhang, Chong Ruan, et al. Deepseek-v3 technical report. *arXiv preprint arXiv:2412.19437*, 2024.

OpenAI. GPT-5.1 instant and GPT-5.1 thinking system card addendum, 2025.

Bart Selman and Carla P Gomes. Hill-climbing search. *Encyclopedia of cognitive science*, 81 (333-335):10, 2006.

Olivier Chapelle and Lihong Li. An empirical evaluation of Thompson sampling. *Advances in neural information processing systems*, 24, 2011.

Shipra Agrawal and Navin Goyal. Thompson sampling for contextual bandits with linear payoffs. In *Proceedings of International Conference on Machine Learning*, pages 127–135. PMLR, 2013.

John Langford and Tong Zhang. The epoch-greedy algorithm for multi-armed bandits with side information. *Advances in neural information processing systems*, 20, 2007.

Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Pearson, 4 edition, 2020.

Grzegorz Dudek. Tournament searching method to feature selection problem. In *Artificial Intelligence and Soft Computing*, volume 6114 of *Lecture Notes in Computer Science*, pages 437–444. Springer, 2010. doi: 10.1007/978-3-642-13232-2_53.

James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(2), 2012.

Tobias Brünner, Jochen Reiner, Martin Natter, and Bernd Skiera. Prospect theory in a dynamic game: Theory and evidence from online pay-per-bid auctions. *Journal of Economic Behavior & Organization*, 164:215–234, 2019.

José M Bernardo. Expected information as expected utility. *the Annals of Statistics*, pages 686–690, 1979.

APPENDIX CONTENTS

A Additional Theoretical Analysis **15**

A.1 Relationship Between **BUFS** and **MUFS** 15

A.1.1 Decomposition of the Objectives of **BUFS** and **MUFS** 15

A.1.2 When Do **BUFS** and **MUFS** Coincide 15

A.2 Proof of Proposition 1 16

A.3 Discussion about Assumption in Proposition 1 and Implementation 19

B Additional Experimental Results **21**

B.1 Posterior Update Dynamics 21

B.2 Robustness to Model-Pool Misspecification 21

B.3 Effect of Model-Pool Size under Misspecification 22

B.4 Evaluation Beyond LLM-Based Agents 22

C Experiment Implementation Details **23**

C.1 Introduction about Tested Auctions 23

C.1.1 The First-Price Auctions 23

C.1.2 The Pay-per-Bid Auctions 23

C.2 Hyperparameter Settings 23

C.2.1 The First-Price Auctions 24

C.2.2 The Pay-per-bid Auctions 24

C.3 Algorithm Details 24

C.3.1 BUFS 24

C.3.2 MUFS 25

C.4 Behavioral Models Proposed by LLMs 26

C.4.1 The First-Price Auctions 26

C.4.2 The Pay-per-Bid Auctions 29

C.5 Implementation Details with the LLMs 31

C.5.1 Prompt Used in the First-Price Auctions 31

C.5.2 Prompt Used in the Pay-per-Bid Auctions 32

A Additional Theoretical Analysis

A.1 Relationship Between **BUFS** and **MUFS**

A.1.1 Decomposition of the Objectives of **BUFS** and **MUFS**

We provide a detailed derivation of the relationship between the objectives of **BUFS** and **MUFS**. Since $\boldsymbol{z}^*$ is a deterministic function of $(\phi, \boldsymbol{\theta})$, the following identity holds (Bernardo, 1979; Zhong et al., 2026)

$$\begin{aligned} &I(\phi, \boldsymbol{\theta}, \boldsymbol{z}^*; \boldsymbol{y}_t \mid \boldsymbol{e}, \mathcal{D}_{t-1}) \\ &\quad = I(\phi, \boldsymbol{\theta}; \boldsymbol{y}_t \mid \boldsymbol{e}, \mathcal{D}_{t-1}) + I(\boldsymbol{z}^*; \boldsymbol{y}_t \mid \phi, \boldsymbol{\theta}, \boldsymbol{e}, \mathcal{D}_{t-1}) \\ &\quad = I(\phi, \boldsymbol{\theta}; \boldsymbol{y}_t \mid \boldsymbol{e}, \mathcal{D}_{t-1}), \end{aligned} \tag{13}$$

Meanwhile, applying the chain rule gives

$$\begin{aligned} &I(\phi, \boldsymbol{\theta}, \boldsymbol{z}^*; \boldsymbol{y}_t \mid \boldsymbol{e}, \mathcal{D}_{t-1}) \\ &\quad = I(\boldsymbol{z}^*; \boldsymbol{y}_t \mid \boldsymbol{e}, \mathcal{D}_{t-1}) + I(\phi, \boldsymbol{\theta}; \boldsymbol{y}_t \mid \boldsymbol{z}^*, \boldsymbol{e}, \mathcal{D}_{t-1}). \end{aligned} \tag{14}$$

Combining (13) and (14) yields

$$I(\phi, \boldsymbol{\theta}; \boldsymbol{y}_t \mid \boldsymbol{e}, \mathcal{D}_{t-1}) = I(\boldsymbol{z}^*; \boldsymbol{y}_t \mid \boldsymbol{e}, \mathcal{D}_{t-1}) + I(\phi, \boldsymbol{\theta}; \boldsymbol{y}_t \mid \boldsymbol{z}^*, \boldsymbol{e}, \mathcal{D}_{t-1}), \tag{15}$$

where the first term on the left is the objective of **BUFS**, and the first term on the right is the objective of **MUFS**. The last term captures information about the model-parameter pair that remains after the induced optimal mechanism parameter is already known. Thus, it represents information that is not directly needed to identify $\boldsymbol{z}^*$.

A.1.2 When Do **BUFS** and **MUFS** Coincide

Equation (15) shows that **BUFS** and **MUFS** coincide when $I((\phi, \boldsymbol{\theta}); \boldsymbol{y}_t \mid \boldsymbol{z}^*, \boldsymbol{e}, \mathcal{D}_{t-1}) = 0$. This condition means that, once the induced optimal mechanism parameter $\boldsymbol{z}^*$ is known, observing $\boldsymbol{y}_t$ provides no additional information about which model-parameter pair $(\phi, \boldsymbol{\theta})$ generated it. We give two sufficient conditions for this condition to hold.

Injective mapping to the optimal mechanism parameter. Suppose the mapping $(\phi, \boldsymbol{\theta}) \mapsto \boldsymbol{z}^*(\phi, \boldsymbol{\theta})$ is injective on the support of the posterior $p(\phi, \boldsymbol{\theta} \mid \mathcal{D}_{t-1})$. Then each value of $\boldsymbol{z}^*$ uniquely determines the corresponding $(\phi, \boldsymbol{\theta})$. Hence, conditional on $\boldsymbol{z}^*$, there is no remaining uncertainty about $(\phi, \boldsymbol{\theta})$, which implies $I(\phi, \boldsymbol{\theta}; \boldsymbol{y}_t \mid \boldsymbol{z}^*, \boldsymbol{e}, \mathcal{D}_{t-1}) = 0$.

Predictive equivalence among model-parameter pairs inducing the same optimal mechanism parameter. For any $\boldsymbol{z}$, define the equivalence class

$$\mathcal{C}_{\boldsymbol{z}} = \{(\phi, \boldsymbol{\theta}) : \boldsymbol{z}^*(\phi, \boldsymbol{\theta}) = \boldsymbol{z}\}.$$

Suppose that, for every candidate experiment $\boldsymbol{e}$, all model-parameter pairs within the same equivalence class induce the same predictive distribution. That is, for any $(\phi_1, \boldsymbol{\theta}_1), (\phi_2, \boldsymbol{\theta}_2) \in \mathcal{C}_{\boldsymbol{z}}$, $p(\boldsymbol{y}_t \mid \boldsymbol{e}, \phi_1, \boldsymbol{\theta}_1, \mathcal{D}_{t-1}) = p(\boldsymbol{y}_t \mid \boldsymbol{e}, \phi_2, \boldsymbol{\theta}_2, \mathcal{D}_{t-1})$. Then, conditional on $\boldsymbol{z}^*$, the predictive distribution of $\boldsymbol{y}_t$ no longer depends on the particular $(\phi, \boldsymbol{\theta})$ within $\mathcal{C}_{\boldsymbol{z}}$. Therefore,

$$\boldsymbol{y}_t \perp\!\!\!\perp (\phi, \boldsymbol{\theta}) \mid \boldsymbol{z}^*, \boldsymbol{e}, \mathcal{D}_{t-1},$$

which again implies $I(\phi, \boldsymbol{\theta}; \boldsymbol{y}_t \mid \boldsymbol{z}^*, \boldsymbol{e}, \mathcal{D}_{t-1}) = 0$.

A.2 Proof of Proposition 1

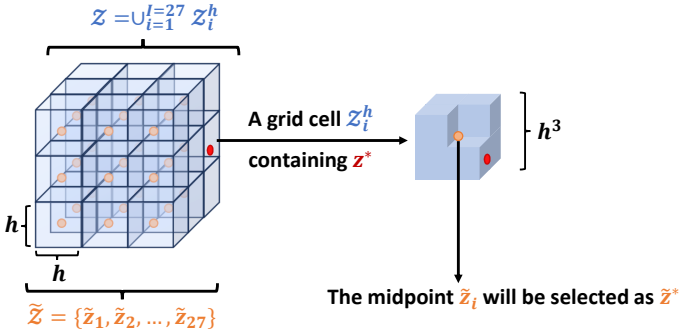


Figure 4: Illustration of the uniform discretization of $\mathcal{Z} \subset \mathbb{R}^3$

Figure 4 illustrates the uniform discretization procedure described in the Proposition 1 for the case where $\mathcal{Z} \subset \mathbb{R}^3$ and the subspace $\tilde{\mathcal{Z}}$ has 27 points. The proof consists of two main steps. First, we decompose the entropy to show that the difference between the posterior entropy induced by $\boldsymbol{z}^*$ and the posterior entropy induced by $\tilde{\boldsymbol{z}}^*$ is governed by a constant term $d \log h$ and a KL divergence term. Second, we show that this KL term admits a uniform bound. When the bound is smaller than the minimum nonzero entropy gap Δ, the discretization of space $\mathcal{Z}$ cannot alter any strict ordering, and thus the ranking is preserved. The detailed proof is provided below.

For simplicity, let

$$p(\boldsymbol{z}^*) := p(\boldsymbol{z}^* \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1}), \qquad \tilde{p}_h(\boldsymbol{z}^*) := \tilde{p}_h(\boldsymbol{z}^* \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1}). \tag{16}$$

Recall that the *probability mass of cell* $\mathcal{Z}_i^h$ is

$$\pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t) = \int_{\mathcal{Z}_i^h} p(\boldsymbol{u} \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1})\, d\boldsymbol{u}, \tag{17}$$

where $p(\boldsymbol{u} \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1})$ denotes the posterior density of $\boldsymbol{z}^*$ and $\boldsymbol{u}$ is the dummy variable. The dependence of $\pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t)$ on $(\boldsymbol{e}, \boldsymbol{y}_t)$ is kept explicit to emphasize its variation across candidate experiments and potential observations, whereas the dependence on $\mathcal{D}_{t-1}$ is suppressed for simplicity.

By the definition of the *piecewise-constant density*, for $\boldsymbol{z}^* \in \mathcal{Z}_i^h$,

$$\tilde{p}_h(\boldsymbol{z}^*) = \frac{\pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t)}{h^d}. \tag{18}$$

Intuitively, $\tilde{p}_h$ *spreads the posterior probability mass* $\pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t)$ *uniformly over the cell* $\mathcal{Z}_i^h$. Thus, it is constant within each cell but preserves the total posterior mass of every cell. For each cell $\mathcal{Z}_i^h$,

$$\int_{\mathcal{Z}_i^h} \tilde{p}_h(\boldsymbol{z})\, d\boldsymbol{z} = \int_{\mathcal{Z}_i^h} p(\boldsymbol{z})\, d\boldsymbol{z} = \pi_i^h, \tag{19}$$

and the density integrates to one over the entire space:

$$\int_{\mathcal{Z}} \tilde{p}_h(\boldsymbol{z})\, d\boldsymbol{z} = \sum_i \int_{\mathcal{Z}_i^h} \tilde{p}_h(\boldsymbol{z})\, d\boldsymbol{z} = \sum_i \pi_i^h = 1. \tag{20}$$

Intuitively, the randomness under $\tilde{p}_h$ can be decomposed into two stages: first selecting a cell according to the masses $\{\pi_i^h\}_{i=1}^I$; second, given the selected cell, sampling uniformly within that cell. The first stage contributes the discrete entropy of the cell probabilities, while the second stage contributes

the within-cell uniform entropy $\log h^d = d \log h$, which is the same for every cell. Therefore,

$$\begin{aligned} H(\tilde{p}_h) &= -\int_{\mathcal{Z}} \tilde{p}_h(\boldsymbol{z}^*) \log \tilde{p}_h(\boldsymbol{z}^*)\, d\boldsymbol{z}^* \\ &= -\sum_{i=1}^{I} \int_{\mathcal{Z}_i^h} \frac{\pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t)}{h^d} \log \frac{\pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t)}{h^d}\, d\boldsymbol{z}^* \\ &= -\sum_{i=1}^{I} \pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t) \log \frac{\pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t)}{h^d} \\ &= -\sum_{i=1}^{I} \pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t) \log \pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t) + \sum_{i=1}^{I} \pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t) \log h^d. \end{aligned} \tag{21}$$

Since $\sum_{i=1}^{I} \pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t) = 1$ and $\log h^d = d \log h$, we have

$$H(\tilde{p}_h) = -\sum_{i=1}^{I} \pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t) \log \pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t) + d \log h. \tag{22}$$

Recall we assume that if $\boldsymbol{z}^* \in \mathcal{Z}_i^h$, then $\tilde{\boldsymbol{z}}^* = \tilde{\boldsymbol{z}}_i$, indicating that $\tilde{\boldsymbol{z}}^*$ takes the midpoint of the cell that contains $\boldsymbol{z}^*$. Since $\pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t)$ is the posterior probability that $\boldsymbol{z}^*$ falls in the i-th cell, it is also the probability that $\tilde{\boldsymbol{z}}^*$ takes the value $\tilde{\boldsymbol{z}}_i$. *Thus, the randomness of $\tilde{\boldsymbol{z}}^*$ is exactly the randomness of which cell contains $\boldsymbol{z}^*$.* It implies

$$P(\tilde{\boldsymbol{z}}^* = \tilde{\boldsymbol{z}}_i \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1}) = P(\boldsymbol{z}^* \in \mathcal{Z}_i^h \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1}) = \pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t). \tag{23}$$

Therefore, the entropy of the discrete distribution $\{\pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t)\}_{i=1}^{I}$ equals the entropy of $\tilde{\boldsymbol{z}}^*$, and we have

$$H(\tilde{p}_h) = H(p(\tilde{\boldsymbol{z}}^* \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1})) + d \log h. \tag{24}$$

Next, we will show the relation between $H(p(\tilde{\boldsymbol{z}}^* \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1}))$ and the KL divergence $D_{\mathrm{KL}}(p \,\|\, \tilde{p}_h)$.

$$\begin{aligned} D_{\mathrm{KL}}(p \,\|\, \tilde{p}_h) &= \int_{\mathcal{Z}} p(\boldsymbol{z}^*) \log \frac{p(\boldsymbol{z}^*)}{\tilde{p}_h(\boldsymbol{z}^*)}\, d\boldsymbol{z}^* \\ &= \int_{\mathcal{Z}} p(\boldsymbol{z}^*) \log p(\boldsymbol{z}^*)\, d\boldsymbol{z}^* - \int_{\mathcal{Z}} p(\boldsymbol{z}^*) \log \tilde{p}_h(\boldsymbol{z}^*)\, d\boldsymbol{z}^*. \end{aligned} \tag{25}$$

Using (18) and (17), the second term in (25) becomes

$$\begin{aligned} \int_{\mathcal{Z}} p(\boldsymbol{z}^*) \log \tilde{p}_h(\boldsymbol{z}^*)\, d\boldsymbol{z}^* &= \sum_{i=1}^{I} \int_{\mathcal{Z}_i^h} p(\boldsymbol{z}^*) \log \frac{\pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t)}{h^d}\, d\boldsymbol{z}^* \\ &= \sum_{i=1}^{I} \pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t) \log \frac{\pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t)}{h^d} && (26) \\ &= \int_{\mathcal{Z}} \tilde{p}_h(\boldsymbol{z}^*) \log \tilde{p}_h(\boldsymbol{z}^*)\, d\boldsymbol{z}^*, && (27) \\ &= -H(\tilde{p}_h), && (28) \end{aligned}$$

where (26) follows because $\log(\pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t)/h^d)$ is constant on $\mathcal{Z}_i^h$ and $\int_{\mathcal{Z}_i^h} p(\boldsymbol{z}^*) d\boldsymbol{z}^* = \pi_i^h(\boldsymbol{e}, \boldsymbol{y}_t)$, and (27) follows because $\tilde{p}_h$ is constant on each cell with value π_i^h/h^d, and each cell has volume h^d. Therefore,

$$D_{\mathrm{KL}}(p \,\|\, \tilde{p}_h) = -H(p(\boldsymbol{z}^* \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1})) + H(\tilde{p}_h). \tag{29}$$

Substituting (24) yields

$$H(p(\boldsymbol{z}^* \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1})) = H(p(\tilde{\boldsymbol{z}}^* \mid \boldsymbol{e}, \boldsymbol{y}_t, \mathcal{D}_{t-1})) + d \log h - D_{\mathrm{KL}}\left(p \,\|\, \tilde{p}_h\right). \tag{30}$$

Next, we will bound the KL divergence term in (30) to control the discrepancy between p and its piecewise-constant approximation $\tilde{p}_h$. This in turn controls the difference between the original posterior entropy $H(p(\boldsymbol{z}^*))$ and the adjusted surrogate entropy $H(p(\tilde{\boldsymbol{z}}^*)) + d \log h$.

For each cell $\mathcal{Z}_i^h$, denote by

$$\bar{p}_i := \frac{1}{h^d} \int_{\mathcal{Z}_i^h} p(\boldsymbol{u})\, d\boldsymbol{u} = \frac{\pi_i^h}{h^d} \tag{31}$$

the average height of the posterior density over this cell. Thus, $\bar{p}_i$ *is exactly the constant value taken by* $\tilde{p}_h$ *on cell* $\mathcal{Z}_i^h$, so we have

$$\tilde{p}_h(\boldsymbol{z}^*) = \bar{p}_i \quad \text{for} \quad \boldsymbol{z}^* \in \mathcal{Z}_i^h. \tag{32}$$

Therefore,

$$D_{\mathrm{KL}}(p \| \tilde{p}_h) = \sum_{i=1}^{I} \int_{\mathcal{Z}_i^h} p(\boldsymbol{z}^*) \log \frac{p(\boldsymbol{z}^*)}{\bar{p}_i}\, d\boldsymbol{z}^*. \tag{33}$$

For each cell, define $\varphi_i(a) := a \log(a/\bar{p}_i)$. Then $\varphi_i(p(\boldsymbol{z}^*)) = p(\boldsymbol{z}^*) \log(p(\boldsymbol{z}^*)/\bar{p}_i)$, $\varphi_i(\bar{p}_i) = 0$, $\varphi_i'(\bar{p}_i) = 1$, and $\varphi_i''(a) = 1/a$. By the assumed lower bound $p(\boldsymbol{z}^*) \geq m > 0$, we also have $\bar{p}_i \geq m > 0$ for every cell, since $\bar{p}_i$ is the average of p over cell $\mathcal{Z}_i^h$. Thus, every point between $p(\boldsymbol{z}^*)$ and $\bar{p}_i$ is also at least m. Therefore, by Taylor's theorem, for some ξ between $p(\boldsymbol{z}^*)$ and $\bar{p}_i$,

$$\begin{aligned} \varphi_i(p(\boldsymbol{z}^*)) &= p(\boldsymbol{z}^*) - \bar{p}_i + \frac{1}{2}\varphi_i''(\xi)\big(p(\boldsymbol{z}^*) - \bar{p}_i\big)^2 \\ &\leq p(\boldsymbol{z}^*) - \bar{p}_i + \frac{1}{2m}\big(p(\boldsymbol{z}^*) - \bar{p}_i\big)^2. \end{aligned} \tag{34}$$

Equivalently,

$$p(\boldsymbol{z}^*) \log \frac{p(\boldsymbol{z}^*)}{\bar{p}_i} \leq p(\boldsymbol{z}^*) - \bar{p}_i + \frac{1}{2m}\big(p(\boldsymbol{z}^*) - \bar{p}_i\big)^2. \tag{35}$$

Integrating over $\mathcal{Z}_i^h$ on both sides, and using $\int_{\mathcal{Z}_i^h} (p(\boldsymbol{z}^*) - \bar{p}_i) d\boldsymbol{z}^* = 0$, we have

$$\int_{\mathcal{Z}_i^h} p(\boldsymbol{z}^*) \log \frac{p(\boldsymbol{z}^*)}{\bar{p}_i} d\boldsymbol{z}^* \leq \frac{1}{2m} \int_{\mathcal{Z}_i^h} (p(\boldsymbol{z}^*) - \bar{p}_i)^2 d\boldsymbol{z}^*. \tag{36}$$

Therefore,

$$D_{\mathrm{KL}}(p \| \tilde{p}_h) \leq \frac{1}{2m} \sum_{i=1}^{I} \int_{\mathcal{Z}_i^h} \big(p(\boldsymbol{z}^*) - \bar{p}_i\big)^2\, d\boldsymbol{z}^*, \tag{37}$$

where $\bar{p}_i$ *is the cell average making the squared error minimized.* Given $\tilde{\boldsymbol{z}}_i$ as the midpoint of cell $\mathcal{Z}_i^h$, we have

$$\int_{\mathcal{Z}_i^h} \big(p(\boldsymbol{z}^*) - \bar{p}_i\big)^2\, d\boldsymbol{z}^* \leq \int_{\mathcal{Z}_i^h} \big(p(\boldsymbol{z}^*) - p(\tilde{\boldsymbol{z}}_i)\big)^2\, d\boldsymbol{z}^*. \tag{38}$$

By assumption, p is G-Lipschitz on each cell. That is, for any $\boldsymbol{z}, \boldsymbol{z}' \in \mathcal{Z}_i^h$,

$$|p(\boldsymbol{z}) - p(\boldsymbol{z}')| \leq G \|\boldsymbol{z} - \boldsymbol{z}'\|_2. \tag{39}$$

Thus,

$$\int_{\mathcal{Z}_i^h} \big(p(\boldsymbol{z}^*) - \bar{p}_i\big)^2 \, d\boldsymbol{z}^* \le \int_{\mathcal{Z}_i^h} \big(p(\boldsymbol{z}^*) - p(\tilde{\boldsymbol{z}}_i)\big)^2 \, d\boldsymbol{z}^* \le G^2 \int_{\mathcal{Z}_i^h} \|\boldsymbol{z}^* - \tilde{\boldsymbol{z}}_i\|_2^2 \, d\boldsymbol{z}^*. \tag{40}$$

For a d-dimensional cube with side length h,

$$\int_{\mathcal{Z}_i^h} \|\boldsymbol{z}^* - \tilde{\boldsymbol{z}}_i\|_2^2 \, d\boldsymbol{z}^* = \frac{d}{12} h^{d+2}. \tag{41}$$

Intuitively, this integral equals the cell volume times the average squared distance from a uniformly chosen point in the cell to its midpoint.

Given (37), we have

$$D_{\mathrm{KL}}(p\|\tilde{p}_h) \le \frac{1}{2m} \sum_{i=1}^{I} G^2 \frac{d}{12} h^{d+2} = \frac{dG^2}{24m} I h^{d+2}. \tag{42}$$

For simplicity, let

$$C_{\mathrm{KL}} := \frac{dG^2 I h^d}{24m}. \tag{43}$$

Then

$$D_{\mathrm{KL}}(p\|\tilde{p}_h) \le C_{\mathrm{KL}} h^2. \tag{44}$$

Then, for each $\boldsymbol{e} \in \mathcal{E}$, we have

$$\tilde{\mathcal{S}}(\boldsymbol{e}) = \mathcal{S}(\boldsymbol{e}) + K_h(\boldsymbol{e}), \tag{45}$$

where

$$K_h(\boldsymbol{e}) := \mathbb{E}_{p(\boldsymbol{y}_t|\boldsymbol{e},\mathcal{D}_{t-1})} \left[D_{\mathrm{KL}}(p\|\tilde{p}_h) \right], \qquad 0 \le K_h(\boldsymbol{e}) \le C_{\mathrm{KL}} h^2. \tag{46}$$

For any $\boldsymbol{e}_1, \boldsymbol{e}_2 \in \mathcal{E}$, if $\mathcal{S}(\boldsymbol{e}_1) < \mathcal{S}(\boldsymbol{e}_2)$, then $\mathcal{S}(\boldsymbol{e}_1) + \Delta \le \mathcal{S}(\boldsymbol{e}_2)$. Hence

$$\tilde{\mathcal{S}}(\boldsymbol{e}_2) - \tilde{\mathcal{S}}(\boldsymbol{e}_1) = \mathcal{S}(\boldsymbol{e}_2) - \mathcal{S}(\boldsymbol{e}_1) + K_h(\boldsymbol{e}_2) - K_h(\boldsymbol{e}_1) \ge \Delta - C_{\mathrm{KL}} h^2. \tag{47}$$

Therefore, if $C_{\mathrm{KL}} h^2 < \Delta$, equivalently

$$h < \sqrt{\frac{24m\Delta}{dG^2 I h^d}}, \tag{48}$$

then $\tilde{\mathcal{S}}(\boldsymbol{e}_1) < \tilde{\mathcal{S}}(\boldsymbol{e}_2)$.

Further, let $A_h := I h^d$ denote the total volume of the grid cells, then we have $h < \sqrt{24m\Delta/(dG^2 A_h)}$. Therefore, the strict pairwise rankings of experiments induced by $\tilde{\boldsymbol{z}}^*$ are preserved under $\boldsymbol{z}^*$.

This completes our proof.

A.3 Discussion about Assumption in Proposition 1 and Implementation

Overall, we acknowledge the gap between the theoretical setting of Proposition 1 and the implemented **MUFS**. For example, the Lipschitz continuity and positive lower-bound assumptions are sufficient rather than necessary conditions for ranking preservation. Therefore, their violation does not by itself imply a failure of the surrogate ranking. However, these assumptions can be difficult to

verify in practice and may not hold throughout sequential updating. In particular, the lower-bound assumption may fail as the posterior over z^* becomes increasingly concentrated.

Empirically, **MUFS** performs well without explicitly verifying these conditions, suggesting that the surrogate space remains effective in practice beyond the sufficient conditions covered by Proposition 1. Therefore, we view Proposition 1 as a continuous-space approximation result that provides sufficient conditions for ranking preservation and theoretical motivation for the surrogate construction, rather than as a guarantee that directly applies to every practical implementation.

B Additional Experimental Results

B.1 Posterior Update Dynamics

Figure 5 visualizes the posterior updates of the five sequential design methods at $t = 1, 2, 5$. For each method, we report the run whose final revenue is closest to the mean value in Table 2, and all runs start from uniform priors. After the first two rounds, **MUFS** and **BUFS** exhibit similar updates of $\boldsymbol{\theta}^{(1)}$ and $\boldsymbol{\theta}^{(2)}$, and by $t = 5$ their beliefs become largely aligned. In contrast, **Seq-Random** and **Seq-HC** follow different update dynamics and retain non-negligible belief in $\phi^{(2)}$. Notably, both **MUFS** and **BUFS** assign nearly zero belief to $\phi^{(2)}$ by $t = 5$, effectively ruling it out as either the model contributing to optimization or the true behavioral model. This contrast reveals that our methods update the posterior more efficiently and eliminate uninformative candidates earlier, thereby allocating the experimental budget more effectively. Although **Seq-Random** and **Seq-HC** improve over their non-sequential counterparts (i.e., **Random Sampling** and **Heuristic Search**), the gap to **BUFS** and **MUFS** highlights the importance of an informative experimental design criterion for mechanism parameter optimization.

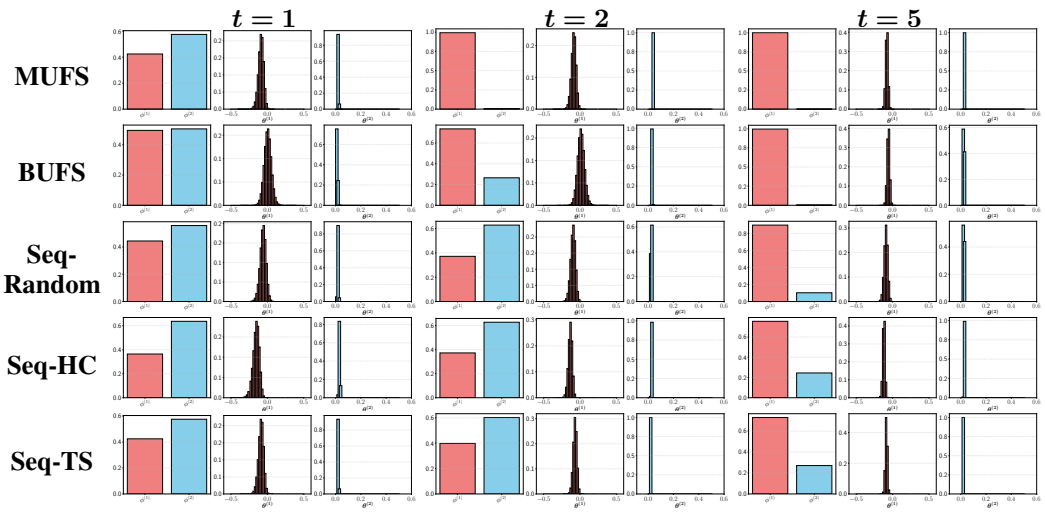


Figure 5: Posterior belief update dynamics over ϕ and $\boldsymbol{\theta}$ under GPT-5.1 in pay-per-bid auctions. Rows correspond to different methods, and columns show the posterior beliefs at $t = 1, 2, 5$. For each selected round, the first panel shows the posterior over $(\phi^{(1)}, \phi^{(2)})$, followed by the discrete posteriors over $\boldsymbol{\theta}^{(1)}$ and $\boldsymbol{\theta}^{(2)}$, respectively.

B.2 Robustness to Model-Pool Misspecification

In our framework, LLMs are used to propose candidate behavioral models that form the hypothesis pool. We do not assume that this pool covers the full diversity of human behavior or contains the exact model governing the behavior of the target agents. Instead, the framework can remain effective when the candidate models only approximate the target behavior and capture some of its behaviorally relevant features.

To directly examine this, we conduct a new controlled experiment using a human-data-calibrated oracle model. Following Brünner et al. (2019), we fit the oracle model to real human behavioral data and treat it as the ground-truth agent interacting with different experimental design methods. We consider two settings, each with a candidate pool of three models. The two settings share the same two LLM-generated models and differ only in the remaining model:

- **Oracle-included setting:** the pool contains the oracle model together with two LLM-generated models and is therefore correctly specified. Note that one of the behavioral parameters of the oracle model remains unknown to the methods and must be inferred from sequential observations.

- **Oracle-excluded setting:** the oracle model is replaced by a third LLM-generated model. Hence, all three candidate models are misspecified relative to the oracle.

Since performance is evaluated against the same oracle agent in both settings, and the candidate pools differ only in whether the oracle model is included, this comparison isolates the effect of model-pool misspecification. Reported results in Table 3 are averaged over 10 independent trials with different random seeds. Specifically, the metrics in Table 3 are defined as follows:

- **Mean**: Average revenue achieved under the optimized mechanism parameter.
- **Std.**: Standard deviation of the revenue across independent trials.
- **Mean Regret**: Difference between the maximum achievable oracle revenue and the mean revenue.
- **Exceedance Rate**: Proportion of feasible mechanism designs with lower expected oracle revenue. For example, 99.64% means that the method's mean revenue is higher than that of 99.64% of the mechanism designs in the feasible design space.

Table 3: Comparison in pay-per-bid auctions under oracle agents. The best values are in **bold**, the second-best values are underlined, and the arrow indicates the preferred direction.

Model Pool Index	Method	Mean ↑	Std. ↓	Mean Regret ↓	Exceedance Rate ↑
Oracle included	**MUFS**	<u>28.08</u>	<u>0.05</u>	<u>1.22</u>	<u>99.64%</u>
Oracle excluded	**MUFS**	27.23	2.45	2.07	99.04%
Oracle included	**BUFS**	**28.11**	**0.00**	**1.19**	**99.67%**
Oracle excluded	**BUFS**	27.20	2.57	2.10	98.99%

As expected, Table 3 shows that including the oracle model yields the best performance. More importantly, even when the hypothesis pool contains only LLM-proposed models, both **MUFS** and **BUFS** still display strong performance in mechanism optimization.

Although the LLM-generated models in the oracle-excluded setting are misspecified, the pool still shares a behaviorally relevant feature with the oracle: one candidate model also accounts for bidders' risk preferences (see Appendix C.4.2). This may help explain why the misspecified pool remains informative for mechanism optimization, consistent with the motivation in Section 4.1 that an expressive candidate pool can provide informative approximations of the target behavior.

B.3 Effect of Model-Pool Size under Misspecification

We follow the controlled experiment in Appendix B.2, but here consider 4 LLM-generated candidate models that are all misspecified relative to the oracle to varying degrees (see Appendix C.4.2). To investigate how the number of misspecified models affects the performance, we vary the number of misspecified models in the pool (i.e., pool size). All tested methods follow Section 5.2. To reduce the influence of any particular pool composition, we exhaustively evaluate all six possible size-two pools and all four possible size-three pools. Each pool composition is tested over 10 independent runs with different random seeds. The metrics in Table 4 are defined in Appendix B.2.

Table 4 shows that **MUFS** and **BUFS** consistently outperform the sequential baselines under both pool sizes. Moreover, their performance improves as the misspecified pool expands from two to three models, suggesting that a richer candidate pool may provide a more informative approximation of the target behavior even when none of the candidate models is correctly specified. This is consistent with the motivation in Section 4.1.

B.4 Evaluation Beyond LLM-Based Agents

Experiments in Section 5 use LLM-based agents as the strategic decision-makers, whereas the controlled experiments in Appendix B.3 use a human-data-calibrated oracle model as the participants. Despite this change, **MUFS** and **BUFS** continue to consistently outperform the sequential baselines under fully misspecified candidate pools. These results indicate that our framework does not rely on the strategic agents themselves being LLM-based.

Table 4: Comparison under fully misspecified model pools of different sizes in pay-per-bid auctions with oracle agents. The best values are in **bold**, the second-best values are underlined, and the arrow indicates the preferred direction.

Misspecified Pool Size	Method	Mean ↑	Std. ↓	Mean Regret ↓	Exceedance Rate ↑
2	**BUFS**	25.26	4.66	4.04	96.28%
2	**MUFS**	25.66	4.39	3.64	96.96%
2	Seq-Random	24.46	4.48	4.84	94.31%
2	Seq-HC	23.87	4.65	5.43	92.60%
2	Seq-TS	25.04	4.76	4.26	95.80%
3	**BUFS**	26.82	3.39	2.48	98.70%
3	**MUFS**	**27.28**	2.66	**2.02**	**99.10%**
3	Seq-Random	24.94	4.15	4.36	95.59%
3	Seq-HC	24.55	4.20	4.75	94.63%
3	Seq-TS	26.09	3.87	3.21	97.72%
-	Random Sampling	21.93	2.73	7.37	83.20%
-	Heuristic Search	20.56	**0.97**	8.74	65.33%
-	Coarse Search	21.56	1.56	7.74	80.25%
-	LinTS	25.08	2.51	4.22	95.83%
-	Epoch Greedy	21.82	2.44	7.48	82.43%
-	SPNE	18.76	1.36	10.54	22.19%

C Experiment Implementation Details

C.1 Introduction about Tested Auctions

C.1.1 The First-Price Auctions

In a first-price auction, the seller sets a reserve price r and announces it to all n symmetric bidders. Each bidder i submits a bid b_i. The bidder with the highest bid that exceeds r wins the item and pays their own bid. If no bid exceeds r, the item remains unsold. The reserve price is anonymous, meaning all bidders observe the same r.

In the classical literature, the agents have symmetric independent private values (IPV). Suppose it is a uniform private value between $[v_{\text{low}}, v_{\text{high}}]$, where v_{low} and v_{high} denote the lower and upper bounds of private values, respectively. r could be less than v_{low}, greater than v_{high}, or between low and high value. In the experiments, we set $r \in [1, 2000]$ and private values follow a uniform distribution $[600, 1200]$. The reserve price and the private values are integers.

C.1.2 The Pay-per-Bid Auctions

A pay-per-bid auction has multiple rounds, indexed by $t \in \mathbb{N}^+$. Initially, all bidders may place a bid. The first bid starts the auction and makes the bidder the initial leader. In each subsequent round, any bidder who is not currently leading may place a bid. This bid starts the next round and makes that bidder the new leader.

The auction starts at a price of $q = 0$. Each bid increases the price by a fixed increment $d \geq 0$. Each bidder whose bid is accepted must also pay a non-refundable bid fee $b > 0$. Both b and d are small values (e.g., \$0.1). The auction ends when no further bids are placed. If the last bid occurs in round t, we define the auction duration as $n = t$. The final leading bidder wins the product and pays the final purchase price q.

In the experiments, pay-per-bid auctions differ in bid fees and bid increments. The item and the market price for this item are known to the agents. We set $b \in [0.1, 9]$, $d \in [0.1, 9]$ and the market price is 30. Both b and d have one decimal place.

C.2 Hyperparameter Settings

For reproducibility and practical implementation, we report the hyperparameters used in the two auction settings below.

C.2.1 THE FIRST-PRICE AUCTIONS

The hyperparameter settings used in the first-price auction experiments are summarized in Table 5.

Table 5: Hyperparameters for the first-price auction experiments in Section 5.1.

Symbol	Description	Value
T	Number of experimental rounds	3
K	Number of candidate behavioral models	2
r	Mechanism parameter, the reserve price	$r \in [1, 2000]$
$\tilde{\mathcal{Z}}$	The surrogate space	$\{1, 200, 400, \ldots, 2000\}$
N	Outer Monte Carlo samples	100
M	Mid Monte Carlo samples	50
J	Inner Monte Carlo samples	100

C.2.2 THE PAY-PER-BID AUCTIONS

The hyperparameter settings used in pay-per-bid auction experiments are summarized in Table 6.

Table 6: Hyperparameters for the pay-per-bid auction experiments in Section 5.2.

Symbol	Description	Value
T	Number of experimental rounds	5
K	Number of candidate behavioral models	2
b	Mechanism parameter, the bid fee	$b \in [0.1, 9.0]$
d	Mechanism parameter, the bid increment	$d \in [0.1, 9.0]$
$\tilde{\mathcal{Z}}$	The surrogate space	$\{0.1, 1, 2, \ldots, 9\} \times \{0.1, 1, 2, \ldots, 9\}$
N	Outer Monte Carlo samples	100
M	Mid Monte Carlo samples	50
J	Inner Monte Carlo samples	100

C.3 ALGORITHM DETAILS

C.3.1 BUFS

Algorithm 1: Behavioral Uncertainty-First Strategy (BUFS)

Require: Candidate behavioral models $\{\phi^{(k)}\}_{k=1}^{K}$ with parameter spaces $\{\Theta^{(k)}\}_{k=1}^{K}$; prior $p(\phi, \boldsymbol{\theta})$; experiment space $\mathcal{E}$; mechanism parameter space $\mathcal{Z}$; experimental budget T; Monte Carlo sample sizes N and M; numerical constant $\varepsilon > 0$.

Ensure: Belief-conditioned optimal mechanism parameter $\hat{\boldsymbol{z}}_T$.

$\mathcal{D}_0 \leftarrow \emptyset$

Initialize $p(\phi, \boldsymbol{\theta} \mid \mathcal{D}_0) \leftarrow p(\phi, \boldsymbol{\theta})$

for $t = 1, \ldots, T$ **do**

  for all candidate experiments e **do**

    $\widehat{U}_t(e) \leftarrow 0$

    for $n = 1, \ldots, N$ **do**

      Sample a model index

$$k_n \sim \text{Categorical}\left(\left\{p(\phi = \phi^{(k)} \mid \mathcal{D}_{t-1})\right\}_{k=1}^{K}\right)$$

      Set $\phi_n \leftarrow \phi^{(k_n)}$ and sample

$$\boldsymbol{\theta}_n \sim p\Big(\boldsymbol{\theta}^{(k_n)} \mid \phi = \phi^{(k_n)}, \mathcal{D}_{t-1}\Big)$$

Sample a prospective outcome

$$y_n \sim p(y \mid e, \phi_n, \boldsymbol{\theta}_n, \mathcal{D}_{t-1})$$

Compute the conditional likelihood

$$L_n \leftarrow p(y_n \mid e, \phi_n, \boldsymbol{\theta}_n, \mathcal{D}_{t-1})$$

$\widehat{p}_n \leftarrow 0$

for $m = 1, \ldots, M$ **do**
  Sample $(\phi_{n,m}, \theta_{n,m}) \sim p(\phi, \theta | \mathcal{D}_{t-1})$
  $\widehat{p}_n \leftarrow \widehat{p}_n + p(y_n | e, \phi_{n,m}, \theta_{n,m}, \mathcal{D}_{t-1})$
  $\widehat{p}_n \leftarrow \max\{\frac{1}{M}\widehat{p}_n, \epsilon\}$
end for
$\widehat{p}_n \leftarrow \frac{1}{M}\widehat{p}_n$

$$\widehat{U}_t(e) \leftarrow \widehat{U}_t(e) + \log \frac{L_n}{\widehat{p}_n}$$

  end for
  $\widehat{U}_t(e) \leftarrow \frac{1}{N}\widehat{U}_t(e)$
end for
Select

$$e_t \leftarrow \arg\max_{e \in \mathcal{E}} \widehat{U}_t(e)$$

Conduct experiment e_t and observe y_t^{obs}

$$\mathcal{D}_t \leftarrow \mathcal{D}_{t-1} \cup \{(e_t, y_t^{\text{obs}})\}$$

Update the posterior via Bayes' rule:

$$p(\phi, \boldsymbol{\theta} \mid \mathcal{D}_t) \propto p(y_t^{\text{obs}} \mid e_t, \phi, \boldsymbol{\theta}) p(\phi, \boldsymbol{\theta} \mid \mathcal{D}_{t-1})$$

end for
Compute

$$\hat{\boldsymbol{z}}_T \leftarrow \arg\max_{\boldsymbol{z} \in \mathcal{Z}} \mathbb{E}_{p(\phi, \boldsymbol{\theta} | \mathcal{D}_T)}[f(\boldsymbol{z}, \phi, \boldsymbol{\theta})]$$

return $\hat{\boldsymbol{z}}_T$

C.3.2 MUFS

Algorithm 2: Mechanism Uncertainty-First Strategy (MUFS)

Require: Candidate behavioral models $\{\phi^{(k)}\}_{k=1}^K$ with parameter spaces $\{\Theta^{(k)}\}_{k=1}^K$; prior $p(\phi, \boldsymbol{\theta})$; experiment space $\mathcal{E}$; mechanism parameter space $\mathcal{Z}$; surrogate space $\widetilde{\mathcal{Z}}$; experimental budget T; Monte Carlo sample sizes N, M, and J; numerical constant $\varepsilon > 0$.
Ensure: Belief-conditioned optimal mechanism parameter $\hat{\boldsymbol{z}}_T$.

$\mathcal{D}_0 \leftarrow \emptyset$
Initialize $p(\phi, \boldsymbol{\theta} \mid \mathcal{D}_0) \leftarrow p(\phi, \boldsymbol{\theta})$
for $t = 1, \ldots, T$ **do**
  for all candidate experiments e **do**
    $\widehat{U}_t(e) \leftarrow 0$
    for $n = 1, \ldots, N$ **do**
      Sample a model index

$$k_n \sim \text{Categorical}\left(\left\{p(\phi = \phi^{(k)} \mid \mathcal{D}_{t-1})\right\}_{k=1}^K\right)$$

      Set $\phi_n \leftarrow \phi^{(k_n)}$ and sample

$$\boldsymbol{\theta}_n \sim p\Big(\boldsymbol{\theta}^{(k_n)} \mid \phi = \phi^{(k_n)}, \mathcal{D}_{t-1}\Big)$$

Sample a prospective outcome $y_n \sim p(y \mid e, \phi_n, \boldsymbol{\theta}_n, \mathcal{D}_{t-1})$
Compute the hypothetical posterior

$$p(\phi, \boldsymbol{\theta} \mid e, y_n, \mathcal{D}_{t-1}) \propto p(y_n \mid e, \phi, \boldsymbol{\theta}, \mathcal{D}_{t-1}) p(\phi, \boldsymbol{\theta} \mid \mathcal{D}_{t-1})$$

$S_n \leftarrow 0$
for $m = 1, \ldots, M$ **do**
  Sample $(\phi_{n,m}, \boldsymbol{\theta}_{n,m}) \sim p(\phi, \boldsymbol{\theta} \mid e, y_n, \mathcal{D}_{t-1})$
  Compute

$$\widetilde{\boldsymbol{z}}^*_{n,m} \leftarrow \arg\max_{\widetilde{\boldsymbol{z}} \in \widetilde{\mathcal{Z}}} f(\widetilde{\boldsymbol{z}}, \phi_{n,m}, \boldsymbol{\theta}_{n,m})$$

  $C_{n,m} \leftarrow 0$
  for $j = 1, \ldots, J$ **do**
    Sample $(\phi_{n,m,j}, \boldsymbol{\theta}_{n,m,j}) \sim p(\phi, \boldsymbol{\theta} \mid e, y_n, \mathcal{D}_{t-1})$
    Compute

$$\widetilde{\boldsymbol{z}}^*_{n,m,j} \leftarrow \arg\max_{\widetilde{\boldsymbol{z}} \in \widetilde{\mathcal{Z}}} f(\widetilde{\boldsymbol{z}}, \phi_{n,m,j}, \boldsymbol{\theta}_{n,m,j})$$

    $C_{n,m} \leftarrow C_{n,m} + \mathbb{I}\big(\widetilde{\boldsymbol{z}}^*_{n,m} = \widetilde{\boldsymbol{z}}^*_{n,m,j}\big)$
  end for
  $\widehat{p}_{n,m} \leftarrow \max\left\{\frac{C_{n,m}}{J}, \varepsilon\right\}$
  $S_n \leftarrow S_n + \log \widehat{p}_{n,m}$
end for
$\widehat{U}_t(e) \leftarrow \widehat{U}_t(e) + \frac{1}{M} S_n$
end for
$\widehat{U}_t(e) \leftarrow \frac{1}{N} \widehat{U}_t(e)$
end for
Select

$$e_t \leftarrow \arg\max_{e \in \mathcal{E}} \widehat{U}_t(e)$$

Conduct experiment e_t and observe y_t^{obs}
$\mathcal{D}_t \leftarrow \mathcal{D}_{t-1} \cup \{(e_t, y_t^{\text{obs}})\}$
Update the posterior via Bayes' rule:

$$p(\phi, \boldsymbol{\theta} \mid \mathcal{D}_t) \propto p(y_t^{\text{obs}} \mid e_t, \phi, \boldsymbol{\theta}) p(\phi, \boldsymbol{\theta} \mid \mathcal{D}_{t-1})$$

end for
Compute

$$\hat{\boldsymbol{z}}_T \leftarrow \arg\max_{\boldsymbol{z} \in \mathcal{Z}} \mathbb{E}_{p(\phi, \boldsymbol{\theta} \mid \mathcal{D}_T)}[f(\boldsymbol{z}, \phi, \boldsymbol{\theta})]$$

return $\hat{\boldsymbol{z}}_T$

C.4 Behavioral Models Proposed by LLMs

C.4.1 The First-Price Auctions

The Model Proposed By ChatGPT-5.2 Thinking This model assumes bidders have uniformly distributed private valuations and, if their valuation is high enough such that $\delta v_i > r$, they shade their bid by randomly choosing a bid between the reserve price r and an upper bound equal to a fraction δ of their valuation (i.e., δv_i).

We consider I symmetric bidders. Each bidder i has a private valuation v_i drawn independently from a uniform distribution on $[v_{\text{low}}, v_{\text{high}}]$. The seller announces a reserve price r. Each bidder submits a bid b_i according to the following behavioral rule:

$$b_i = \begin{cases} 0, & \text{if } v_i < \dfrac{r}{\delta}, \\ U_i, & \text{if } v_i \geq \dfrac{r}{\delta}, \end{cases} \tag{49}$$

where U_i is a random variable drawn uniformly from the interval $[r, \delta v_i]$, and $\delta \in (0, 1]$ is a behavioral parameter that captures the average fraction of the valuation that determines the upper bound of the bid. If $\delta v_i < r$, the condition $v_i \geq r/\delta$ fails, so the bid is 0. The bids are independent across bidders conditionally on their valuations.

Distribution of the Highest Bid

Let

$$b = \max_{i=1,\dots,I} b_i$$

denote the highest bid. Its cumulative distribution function (CDF) is given by:

$$F_b(x) = \begin{cases} 0, & x < 0, \\ [p_0]^I, & 0 \leq x < r, \\ [F_{\text{single}}(x)]^I, & r \leq x \leq \delta v_{\text{high}}, \\ 1, & x > \delta v_{\text{high}}, \end{cases} \tag{50}$$

where

$$p_0 = \mathbb{P}\left(v_i < \frac{r}{\delta}\right) = \begin{cases} 0, & \frac{r}{\delta} \leq v_{\text{low}}, \\ \dfrac{r/\delta - v_{\text{low}}}{v_{\text{high}} - v_{\text{low}}}, & v_{\text{low}} < \dfrac{r}{\delta} < v_{\text{high}}, \\ 1, & \dfrac{r}{\delta} \geq v_{\text{high}}, \end{cases} \tag{51}$$

$$F_{\text{single}}(x) = \mathbb{P}(b_i \leq x) = p_0 + \int_r^x f_b(t)\, dt, \quad \text{for } x \in [r, \delta v_{\text{high}}]. \tag{52}$$

The continuous part of the unconditional bid distribution has density

$$f_b(t) = \frac{1}{v_{\text{high}} - v_{\text{low}}} \int_{\max\{v_{\text{low}}, t/\delta\}}^{v_{\text{high}}} \frac{1}{\delta v - r}\, dv, \quad t \in [r, \delta v_{\text{high}}]. \tag{53}$$

Special Case: Uniform Values on $[0, 1]$

For simplicity, assume $v_{\text{low}} = 0$ and $v_{\text{high}} = 1$. Then:

$$p_0 = \begin{cases} \dfrac{r}{\delta}, & 0 \leq r \leq \delta, \\ 1, & r > \delta, \end{cases} \tag{54}$$

$$f_b(t) = \begin{cases} \dfrac{1}{\delta}\left[\ln\left(\dfrac{\delta - r}{t - r}\right)\right], & r \leq t \leq \delta, \\ 0, & \text{otherwise}, \end{cases} \tag{55}$$

$$F_{\text{single}}(x) = p_0 + \frac{x - r}{\delta}\left[1 + \ln\left(\frac{\delta - r}{x - r}\right)\right], \quad r \leq x \leq \delta. \tag{56}$$

Thus, the CDF of the highest bid becomes:

$$F_b(x) = \begin{cases} 0, & x < 0, \\ \left(\dfrac{r}{\delta}\right)^I, & 0 \leq x < r, \\ \left[p_0 + \dfrac{x - r}{\delta}\left(1 + \ln\left(\dfrac{\delta - r}{x - r}\right)\right)\right]^I, & r \leq x \leq \delta, \\ 1, & x > \delta. \end{cases} \tag{57}$$

Behavioral Parameter δ

The parameter $\delta \in (0, 1]$ represents **the degree to which bidders shade their bids relative to their valuations**. A value of $\delta = 1$ means bidders bid uniformly between r and their valuation; $\delta < 1$ implies more aggressive shading (lower bids). In equilibrium for a first-price auction with uniform values and no reserve, the optimal bidding function is linear with slope $(I-1)/I$. Thus, a plausible reference range for δ is around $(I-1)/I$, but it may vary with experience, auction format, or context. Estimation of δ from real bid data can reveal systematic deviations from equilibrium bidding, such as overbidding (δ near 1) or underbidding (δ near 0).

Expected Revenue

The seller's expected revenue is the expectation of the highest bid b:

$$R(r) = \mathbb{E}[b] = \int (1 - F_b(x))\, dx. \tag{58}$$

This integral can be evaluated numerically given the CDF above.

The Model Proposed By Claude Opus 4.5 The behavioral model assumes that the I bidders have independent private valuations v_i drawn from a LogNormal distribution with parameters $\mu = 6.8$ and $\sigma = 0.3$, reflecting a mean valuation of approximately \$940 for the item (calculated as $\exp(\mu + \sigma^2/2) \approx 940$), which is reasonable given the item's description.

Let $F(v)$ denote the CDF of a single bidder's valuation, i.e.,

$$F(v) = \Phi\left(\frac{\ln v - \mu}{\sigma}\right),$$

where Φ is the standard normal CDF, and let $f(v)$ be the corresponding density

$$f(v) = \frac{1}{v\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln v - \mu)^2}{2\sigma^2}\right).$$

The maximum valuation

$$\bar{v} = \max_{i=1,\ldots,I} v_i$$

then has CDF

$$\mathcal{F}_{\bar{V}}(\bar{v}) = [F(\bar{v})]^I$$

and density

$$f_{\bar{V}}(\bar{v}) = I[F(\bar{v})]^{I-1} f(\bar{v}).$$

Behaviorally, we assume bidders exhibit a linear shading bias parameterized by $\delta \in [0.01, 1]$, where δ represents bidding aggressiveness. A high δ means less shading, closer to bidding true value. A low δ means more conservative bidding, closer to reserve.

For all $\bar{v} > 0$, the highest bid is:

$$b = \delta\bar{v} + (1 - \delta)r. \tag{59}$$

This gives:

- If $\bar{v} < r$, then $b < r$. The item is not sold, but the bid is observed.
- If $\bar{v} \geq r$, then $b \geq r$. The item is sold.

The minimum possible bid is

$$b_{\min} = (1 - \delta)r$$

(when $\bar{v} \to 0$).

The **unconditional** density of the highest bid, denoted by $g(b)$, for all $b > (1 - \delta)r$ is:

$$g(b) = \frac{1}{\delta} \cdot f_{\bar{V}}\left(\frac{b - (1 - \delta)r}{\delta}\right). \tag{60}$$

C.4.2 THE PAY-PER-BID AUCTIONS

A game theoretical model proposed by ChatGPT o4-mini Suppose that everyone follows a CARA (constant absolute risk aversion) utility

$$\mu(x) = \frac{1 - e^{-\theta x}}{\theta}.$$

Let $q_t = dt$ be the item price at the t-th round, and let the market value of this item be v. Then the probability that at least one bidder bids in round t is u_t, where

$$\underbrace{(1-u_t)\frac{1-e^{-\theta(v-q_{t-1}-b)}}{\theta}}_{\text{Bid \& win in the } (t-1)\text{-th round}} + \underbrace{u_t\frac{1-e^{-\theta(-b)}}{\theta}}_{\text{Bid \& lose in the } (t-1)\text{-th round}} = \underbrace{0}_{\text{No bid}}, \qquad 1 < t \le T = \left\lfloor \frac{v-b}{d} \right\rfloor. \tag{61}$$

Specifically, $u_t = 1$ when $t = 1$. Then, the probability that the auction ends with duration n is

$$p_n = (1 - u_{n+1}) \prod_{t=1}^{n} u_t. \tag{62}$$

In this model, θ is the behavioral parameter.

A Non-equilibrium model proposed by ChatGPT o4-mini Suppose that each non-leader bids with probability w_t:

$$w_t = \min\{1, \lambda(v - q_t - b)\}, \tag{63}$$

where λ represents the "aggressiveness", and $(v - q_t - b)$ is the remaining surplus that the winner can get in round t.

The probability that at least one bidder bids in round t is

$$u_t = 1 - (1 - w_t)^{I-1}, \qquad 1 < t \le T = \left\lfloor \frac{v-b}{d} \right\rfloor. \tag{64}$$

Specifically, $u_t = 1$ when $t = 1$. Then, the probability that the auction ends with distribution n is

$$p_n = (1 - u_{n+1}) \prod_{t=1}^{n} u_t. \tag{65}$$

A game theoretical model with sunk-cost effects proposed by ChatGPT-5.2 Thinking Suppose that bidders' decisions are affected by previously incurred bidding costs. Let $q_t = dt$ be the item price at the t-th round, let b be the bidding fee, and let the market value of the item be v. After reaching round t, a bidder has already incurred $(t-2)b$ in bidding costs before deciding whether to bid again. We model the perceived sunk-cost effect as

$$s_t(\theta) = \begin{cases} 0, & t = 2, \\ \big((t-2)b\big)^{\theta}, & t > 2, \end{cases} \tag{66}$$

where θ controls the strength of the sunk-cost effect.

Then the probability that at least one bidder bids in round t is u_t, determined by

$$\underbrace{(1-u_t)\left[v-q_{t-1}-b+s_t(\theta)\right]}_{\text{Bid \& win in the } (t-1)\text{-th round}}+\underbrace{u_t\left[-b+s_t(\theta)\right]}_{\text{Bid \& lose in the } (t-1)\text{-th round}}=\underbrace{0}_{\text{No bid}},\qquad 1<t\leq T=\left\lfloor\frac{v-b}{d}\right\rfloor. \tag{67}$$

Solving (67) gives

$$\tilde{u}_t=\frac{v-q_{t-1}-b+s_t(\theta)}{v-q_{t-1}} \tag{68}$$

$$=1+\frac{-b+s_t(\theta)}{v-q_{t-1}}. \tag{69}$$

To guarantee $\tilde{u}_t$ to lie in $[0,1]$, we map it to a valid probability, consistently with our implementation:

$$u_t=\mathrm{clip}_{[0,1]}(\tilde{u}_t)=\min\left\{1,\max\left\{0,\tilde{u}_t\right\}\right\}. \tag{70}$$

Specifically, $u_1=1$ and $u_{T+1}=0$. Then, the probability that the auction ends with distribution n is

$$p_n=(1-u_{n+1})\prod_{t=1}^{n}u_t. \tag{71}$$

In this model, θ is the behavioral parameter governing how previously incurred bidding costs affect bidders' willingness to continue bidding.

This model is only used in experiments in Appendix B.2 and B.3.

A behavioral model with loss aversion and quantal response proposed by Claude Opus 4.5 Suppose that bidders are loss averse and make stochastic bidding decisions according to a quantal-response rule. Let $q_t=td$ be the item price at the t-th round, let b be the bidding fee, let v be the market value of the item, and let I be the number of bidders. We assume that a bidder's subjective probability of winning after bidding in round t is

$$\pi_t=\frac{1}{I-1}. \tag{72}$$

If the bidder wins, the payoff is $v-td-tb$, whereas losing incurs the perceived loss $-\lambda tb$, where $\lambda>1$ is the loss-aversion coefficient. Therefore, the expected utility of bidding in round t is

$$U_t=\pi_t\left(v-td-tb\right)+\left(1-\pi_t\right)\left(-\lambda tb\right). \tag{73}$$

Instead of assuming that bidders deterministically bid whenever $U_t>0$, we adopt a quantal-response specification. The probability that an individual bidder submits a bid in round t is

$$\rho_t=\frac{1}{1+\exp(-\theta U_t)}, \tag{74}$$

where $\theta>0$ controls the sensitivity of bidding decisions to expected utility. Assuming that the remaining $I-1$ bidders make their bidding decisions independently, the probability that at least one bidder bids in round t is

$$u_t = 1 - (1 - \rho_t)^{I-1}, \qquad 1 < t \leq T = \left\lfloor \frac{v}{d+b} \right\rfloor. \tag{75}$$

Specifically, $u_1 = 1$. For $t > T$, we set u_t to approximately zero. Then, the probability that the auction ends with duration n is

$$p_n = (1 - u_{n+1}) \prod_{t=1}^{n} u_t. \tag{76}$$

In this model, θ is the behavioral parameter governing the degree of responsiveness of bidders to their expected utility, while the loss-aversion coefficient λ is fixed.

This model is only used in experiments in Appendix B.3.

C.5 Implementation Details with the LLMs

In each game, all agents receive the same game description, payoff structure, decision space, and instructions about what to do (i.e., what to reply). The only differences are their private values in the first-price auction and their bidding histories in the pay-per-bid auctions.

In the proposed theoretical framework, we follow a standard assumption in the mechanism design literature by assuming agents share a common unknown behavioral model-parameter pair (ϕ, θ), although heterogeneous agents are more realistic in practice. However, we do not prescribe specific values of (ϕ, θ) nor impose any homogeneity constraints on the LLM-based agents in experiments.

All agents are assigned the same personality: "Be neutral and rational. Compare potential gains against losses and risks. Try to make rational decisions." Each agent can see its own name in the prompt, and the names are assigned in alphabetical order, such as Alice, Bob, Cathy, Dylan, etc.

Agents do not retain memory across independent games. Specifically, in the pay-per-bid auction, each bidder is given memory within the same auction so that it can observe its own previous bidding decisions. Because we assume in the framework that, conditional on $(e_t, \phi, \boldsymbol{\theta})$, the outcome y_t is independent of historical observations from previous experiments.

For DeepSeek-V3.1, we use Chinese prompts, while for GPT-5.1, we use English prompts. This choice is based on the languages in which the models are primarily trained. The prompts for the two LLMs differ only in the language used. The Chinese prompts are included in the code zip file, and the English prompts are shown below.

All experiments were conducted on a computer equipped with a 13th Gen Intel(R) Core(TM) i7-13700K processor (16 cores) and an NVIDIA GeForce RTX 4070 GPU with 12GB of VRAM. The LLMs are called through APIs.

C.5.1 Prompt Used in the First-Price Auctions

Prompt for the Agent Bidder

Decision Task ## Please decide the amount you will bid in this auction.
Important Notes

- The reserve price is the minimum threshold to win, but it is not necessarily your optimal bid.
- If you win, your payoff = your private value ${self.private_value_sampled} - your bid.
- Key constraint: To ensure rational decision-making, your bid must be less than or equal to your private value ${self.private_value_sampled}. Bidding above your private value will lead to negative payoff, i.e., a loss.
- Your bidding decision involves three key trade-offs:
  1. **Reserve price**: If your private value is only slightly above the reserve price, bidding close to the reserve price may be advantageous.

2. **Win probability vs. payoff**: A lower bid provides a higher potential payoff but may reduce the chance of winning; a higher bid increases the win probability but may reduce payoff or cause a loss.
3. **Competitor behaviour**: Other bidders' values are also drawn from the same uniform distribution, but you do not know their specific values.

- Your bid must be a specific number.

Output Requirements

- Please output only a number representing your bid amount (in yuan), without any unit.
- **Do not** add any other text, symbols, or explanations.

Your Bid Decision

Prompt for generating the behavioral model

Please propose a reasonable, tractable behavioral model satisfying the following constraints:

1. For a given item, the highest bid among all bids, b, is not unchanged but maybe follow some certain distribution. Therefore, the behavioral model you propose MUST indicate the probability of b. Note that the model has to deal with the case that, when the reserve price is greater than all bids (including b), b is still the highest bid even though now the item will not be sold.
2. The model you propose must contain one (or more, if necessary) parameter(s) worth estimating. It can be behavioral parameters indicating the private value, the preference, decision-making bias, or bounded rationality of agents. Or it can be other parameters related to private information. If there is more than one behavioral parameter, please set the one that is most valuable to estimate with real data as δ and tell me its plausible region. Meanwhile, please provide the commonly observed values of the left behavioral parameters.
3. The behavioral model has to represent probability of $bound[0] \leq b \leq bound[1]$. b can be a positive integer or a positive continuous float. We define the bound as $[0, 3000]$, where 3000 is $\gg$ maximum reserve price.
4. The model DOES NOT have to be an equilibrium model. Any model is welcome. You are also free to add additional assumptions if necessary.
5. You can assume that each agent has her own private valuation $v_i \sim f$ for this item where f is a distribution.

An original abstract visual artwork, printed in a limited run of 100 pieces. Details are as follows:

- Unique edition number: 25/100 (printed on the artwork)
- Medium: High-quality archival print
- Size: 40 $\times$ 60 cm (unframed)
- The image is abstract and does not depict identifiable objects
- Each edition number is printed directly on the artwork
- There is no secondary market for this artwork.
- The artwork is intended solely for personal enjoyment or display.

C.5.2 Prompt Used in the Pay-per-Bid Auctions

Prompt for the Agent Bidder

Pay-per-Bid Auction Rules
Pay-per-Bid auction is a special multi-round ascending auction with these rules:

1. **Bidding:** Each round, participants may bid or not. Among those who bid, one is randomly selected as that round's **temporary leader**.
2. **Leader Fee:** The temporary leader must immediately pay a fixed bid fee of $auction.b to the auctioneer, regardless of the final outcome. Those who do not bid or are not selected pay nothing.
3. **Price Increment:** Each round increases the item price by a fixed rise of $auction.d.

4. **Bidding Restrictions:** The temporary leader of a round cannot bid in the next round. All other participants may choose to bid or not.
5. **End & Win:** The auction ends when no one bids. Since the starting price is $0, the last round's temporary leader wins and pays (Total rounds × $auction.d) for the item.

{auction_status_prompt}
Your Bidding History ## {mem_desc}
Auction History ## Your Name: {self.id}. {history_desc}
Decision Framework ## To maximize your expected profit, you may consider:

1. **Past Costs:** Bid fees already paid as leader.
2. **If You Bid This Round:**
   - You may be selected as temporary leader (random selection in case of a tie)
   - If selected, you pay $auction.b immediately and cannot bid next round
3. **Winning:**
   - You win if you are the leader in a round and no one bids next round.
   - Winner pays the item price: (total number of rounds) × $auction.d.
   - Winner's total profit = market value $auction.v - item price - all bid fees you've paid
4. **Decision Guidance:**
   - Early rounds often present good opportunities (price far below $auction.v)
   - Consider both immediate profit and potential future potential profit
   - Auction may continue many rounds before reaching market value

Your Decision Task

- Based on the above, decide whether to bid this round to maximize your benefits.
- **Output only "Yes" to bid or "No" to not bid. No other text, symbols or explanations.**

Your Decision

Prompt for generating the behavioral model

As a game theorist, you need to model the strategic behavior of rational participants in a "Pay-per-Bid auction" (a special type of ascending auction). The auction rules are described as follows:

- All bidders know the item type (denoted as c) and its market value (denoted as v). To begin with, all bidders can bid. The first bid will make the auction enter its first round. Whoever places the first bid will become the leading bidder in this round. After the first round, any non-leading bidder in the t-th round can bid, triggering the start of the (t + 1)-th round and resulting in the bidder becoming the new leading bidder. If no one bids in the (t + 1)-th round, the auction will end. The leading bidder in the t-th round will be the final winner and the auction's duration n is t.
- The item will start from the buying price q = $0. If multiple people bid simultaneously, the system randomly selects one winner for that round and immediately reveals the identity of the successful bidder. The leading bidder in each round has to pay a fixed bidding cost b (e.g., $0.1). Those who don't bid or whose bids are unsuccessful do not need to pay this fee. Each bid from the leading bidder will raise the buying price q by a fixed amount d (e.g., $0.1). The final winner can buy the item at the final price nd

Please propose a reasonable, tractable behavioral model satisfying the following constraints:

1. The probability $p(n)$ that the auction lasts for n rounds is not unchanged but maybe follow some distribution. If $p(t)$ is the probability that each homogeneous agent bid at round t, u_t is the probability that at least one bidder bids in round t, we have $p_n = (1-u_n+1)\prod_{t=1}^{n} (u_t)$. Therefore, the behavioral model you propose MUST indicate how to calculate the probability of $p(n)$
2. The model does not have to be an equilibrium model. However, the model you propose must contain one (or more, if necessary) behavioral parameter(s) indicating the preference, decision-making bias, or bounded rationality of agents.
3. (a) Given actual auction duration $\hat n$ observed under the game featured by certain b, d, v, c, we can estimate the value of behavioral parameter(s) by standard methods such as MLE.

(b) Given values of b, d, v, c, and behavioral parameter(s), the behavioral model can be seen as a data-generating process of n

(c) If there is more than one behavioral parameter, please set the one that is most valuable to estimate with real data as δ and tell me its plausible region. Meanwhile, please provide the commonly observed values of the remaining behavioral parameters.

4. The model DOES NOT have to be an equilibrium model. Any model is welcome. You are also free to add additional assumptions if necessary.