Sponsored Questions and How to Auction Them

Kshipra Bhawalkar
Google Research
kshipra@google.com

Alexandros Psomas
Purdue University & Google Research
apsomas@purdue.edu

Di Wang
Google Research
wadi@google.com

Abstract

Online platforms connect users with relevant products and services using ads. A key challenge is that a user's search query often leaves their true intent ambiguous. Typically, platforms passively predict relevance based on available signals and in some cases offer query refinements. The shift from traditional search to conversational AI provides a new approach. When a user's query is ambiguous, a Large Language Model (LLM) can proactively offer several clarifying follow-up prompts. In this paper we consider the following: what if some of these follow-up prompts can be "sponsored," i.e., selected for their advertising potential. How should these "suggestion slots" be allocated? And, how does this new mechanism interact with the traditional ad auction that might follow?

This paper introduces a formal model for designing and analyzing these interactive platforms. We use this model to investigate a critical engineering choice: whether it is better to jointly optimize the user interaction and the final ad auction, or to decouple them into separate mechanisms for the suggestion slots and another for the subsequent ad slot. We show that the VCG mechanism can be adopted to jointly optimize the sponsored suggestion and the ads that follow; while this mechanism is more complex, it achieves efficient and truthful outcomes. On the other hand, we prove that the simple-to-implement modular approach suffers from strategic inefficiency: its Price of Anarchy is unbounded, for various natural choices for the two mechanisms. While we show that this inefficiency can be mitigated through revenue redistribution, we argue that the resulting mechanism sacrifices the practical simplicity that motivated the modular design in the first place.

1 Introduction

The rise of generative AI is fundamentally changing how users find information. Instead of a list of links, knowledge panels, etc., users now receive direct answers from Large Language Models (LLMs), followed by suggestions designed to guide their next step. This conversational interface is a powerful tool for resolving a major challenge in ad auctions: ambiguous user intent. For example, a user searching for "running shoes" might be looking for advanced trail running shoes or beginner road running shoes. While there are existing paradigms aiming to address this ambiguity in contemporary search, e.g., related search functionalities [Goo25], suggested chip refinements [Goo24], etc., the shift to direct answers within a conversational flow makes resolving this intent even more pressing. An LLM that offers distinct follow-up prompts or suggestions ("I would like more information about trail running shoes") or direct questions ("Are you a beginner runner?"), can help clarify the user's needs in a single click. This not only improves the user experience by avoiding an overwhelming list of irrelevant information, but also creates a valuable new opportunity for advertisers: by sponsoring a suggestion like "I want to know more about trail running shoes," an advertiser

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

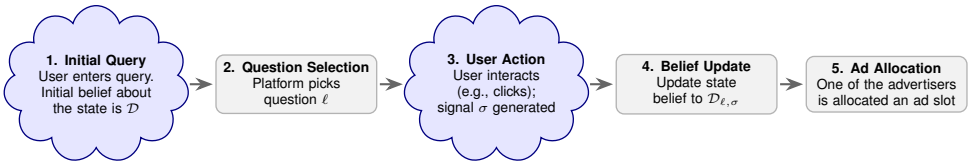

Figure 1: Flowchart of the interactive sponsored search model.

can guide a user with relevant intent directly toward their products, and produce better outcomes for the users, advertisers and platform.[1]

Given the limited display space and the varying utility advertisers derive from different suggestions, the platform faces a critical question: How should these "suggestion" slots be allocated? And, how does this new mechanism for allocating suggestions interact with the traditional ad auction that might follow? In this paper, we introduce a formal model for designing and analyzing these problems. We use this model to investigate a key design choice: whether to build an end-to-end mechanism that jointly optimizes the suggestion and downstream ad auction, or to adopt a simpler modular approach that decouples them into separate auctions.

1.1 Our contributions

We introduce a new model for thinking about advertising in interactive platforms that combines information elicitation and (ad) allocation. While directly inspired by modern applications like sponsored suggestions in conversational AI, a key strength of our model is its generality; our framework is flexible enough to model many different forms of platform-user interaction, including sponsored clickable suggestions (prompts), sponsored questions (that the LLM makes), or other formats.

Our setup is as follows. A platform has a single ad slot for sale to one of n strategic advertisers with quasilinear preferences. The user's intent is captured by a state of nature θ, that encapsulates complex, noisy user features and intent vectors that traditional keyword matching often misses. For the advertiser, θ represents the latent information about the user that is critical for determining an ad's relevance. We assume θ is unknown to all parties and is sampled from a public prior distribution $\mathcal{D}$. Each advertiser i has (1) a private *base* value $v_i \geq 0$, the same across all states, and (2) a *state-dependent* conversion rate $\alpha_i(\theta) \in \mathbb{R}_+$ that represents the ad's relevance for advertiser i given θ.[2]

At the heart of our model is an *information elicitation stage*: alongside any organic content presented to the user, the platform can choose to ask one of k possible sponsored questions. Asking question ℓ generates a signal σ, observable by everyone, drawn from a distribution $Q_\ell(. \mid \theta)$. Here a "question" is the LLM-generated prompt ("I am looking for trail running shoes"), and a "signal" is the publicly observable, measurable outcome of the user clicking or ignoring it. The probability of a user's response depends on their latent intent, captured by the probability distribution $Q_\ell(. \mid \theta)$. Given a signal σ generated by question ℓ, advertisers update their belief about the state θ to a common posterior distribution $\mathcal{D}_{\ell,\sigma}$. This, in turn, implies updated (expected) conversion rates for the ad slot.

Example 1.1 (Running shoes)**.** *A user is asking for information about "running shoes." The platform is uncertain about their specific needs. We model the user's intent as the state of nature* $\theta = (\theta_{terrain}, \theta_{exp}) \in \{0,1\}^2$*, where* $\theta_{terrain} = 1$ *if the user is a* trail *runner and* $\theta_{terrain} = 0$ *if they are a* road *runner;* $\theta_{exp} = 1$ *if the user is an* experienced *runner, and* $\theta_{exp} = 0$ *if they are a* beginner*. The prior belief* $\mathcal{D}$ *is that all four states in* $\Theta = \{(0,0),(0,1),(1,0),(1,1)\}$ *are equally likely; that is,* $\mathcal{D}(\theta) = 1/4$ *for all* $\theta \in \Theta$*.*

There are two advertisers. Advertiser 1 *sells professional trail running shoes. Their private base value is high,* $v_1 = \$50$*. Their conversion rate* $\alpha_1(\theta)$ *is highest for an experienced trail runner, e.g.,* $\alpha_1(1,1) = 0.9$*. It is lower for a beginner trail runner,* $\alpha_1(1,0) = 0.3$*, and zero for all road runners (*$\alpha_1(0,\cdot) = 0$*).* Advertiser 2 *sells popular road running shoes for beginners. Their base value is* $v_2 = \$30$*. Their conversion rate* $\alpha_2(\theta)$ *is highest for a beginner road runner,* $\alpha_2(0,0) = 0.8$*. It is lower for an experienced road runner,* $\alpha_2(0,1) = 0.4$*, and zero for all trail runners (*$\alpha_2(1,\cdot) = 0$*).*

[1] Allowing advertiser interest to steer user conversation should be done with utmost care to ensure high user quality. This paper should be viewed as a theoretical investigation of what such a design would look like; we do not make any claims about feasibility from a policy or regulatory perspective.

[2] This can be, e.g., a click-through-rate, or a conversion value.

Before the ad auction, the platform can show one of two possible sponsored suggestions. One possible choice is ℓ_{trail}: "I am looking for trail running shoes." If the user clicks on this suggestion, this generates a signal $\sigma_{trail,click}$, which, for simplicity, we assume perfectly reveals their interest in trails; that is, if the signal is $\sigma_{trail,click}$, everyone learns that $\theta_{terrain} = 1$. The posterior distribution $\mathcal{D}_{\ell_{trail},\sigma_{trail,click}}$ is the uniform distribution over $(1,0)$ (trail, beginner) and $(1,1)$ (trail, experienced). If the user does not click on this suggestion, this generates a signal $\sigma_{trail,no\ click}$, which, for simplicity, we assume perfectly reveals the user's lack of interest in trails. $\mathcal{D}_{\ell_{trail},\sigma_{trail,no\ click}}$ is the uniform distribution over $(0,0)$ (road, beginner) and $(0,1)$ (road, experienced). A different choice for a suggestion is $\ell_{targeted}$: "I am an experienced trail runner." If clicked, we assume it completely reveals that the state is $\theta = (1,1)$; otherwise, it is revealed that the state is not $\theta = (1,1)$, yielding a uniform posterior distribution over $(0,0)$ (road, beginner), $(1,0)$ (trail, beginner), and $(0,1)$ (road, experienced).

In this paper, we are interested in understanding the trade-offs between two fundamental design philosophies: an end-to-end approach that jointly optimizes the suggestion and auction stages, versus a modular approach that decouples them. Our contributions are as follows:

- We introduce a formal model for sponsored suggestions in interactive platforms.
- We observe that an end-to-end VCG mechanism is truthful (DSIC) in this setting and provide a simple characterization of its payments.
- We consider a natural alternative to the end-to-end approach: a modular mechanism that decouples the auction into two stages (one for the question, and one for the ad). We identify and characterize an intuitive pure Nash equilibrium for the natural "VCG-per-stage" modular mechanism.
- We prove that this modular mechanism has an unbounded Price of Anarchy, and show this persists even when the platform uses first-price or all-pay auctions in the suggestion stage.
- We demonstrate that the modular inefficiency is not inevitable: we prove that a modular mechanism *with redistribution* [Bai97, Cav06] achieves a constant Price of Anarchy, albeit with increased practical complexity.

1.2 Related Work

Many search engines and other products show query refinements or suggested queries to help users. For most popular search engines, there are policies which commit to showing the best results for the user, and which preclude them from considering commercial interest (such as advertiser bids or ad revenue) to optimize these queries [Mic25, Goob]. Exceptions exist; for example, Google's Relevant Search for Content (RSOC), allows publishers to use "funnel RPM" in optimization of suggested queries shown on their webpage [Goo25], thus allowing advertisers to inform which suggestions are shown to the user (note though that Google AFS policies [Gooa] still require that first and foremost user experience is optimized).

Closer to our interest here, there is growing interest from practice to integrate advertising into conversational search. For instance, Perplexity recently launched an experiment featuring sponsored follow-up prompts [Per24]. At the same time, the academic literature on the topic is rapidly growing. Feizi et al. [FHRS25] discuss challenges and opportunities in this area. One major stream of research focuses on how to natively integrate ads into a final, LLM-generated answer. Various models have been proposed, including token-based auctions [DMPL$^+$24], bidding for placement within an LLM summary [DFK$^+$24], fusing organic and sponsored content [MTK24], embedding ads in an independently generated organic answer [HLRS24], and bidding to influence the fine-tuning of the LLM itself [SCS25]. Banchio et al. [BMP25] study the problem of ad placement in a conversational search; specifically, they study equilibria under first- and second-price auctions with respect to the timing of showing an ad. In contrast, our work models the preceding, interactive stage as well: using suggestions to influence the *flow* of the conversation itself. Bergemann et al. [BBD$^+$25] study a setting where agents have private information about both their preferences and a common state, showing that standard mechanisms fail. They propose a solution via "data-driven mechanisms," which use post-allocation data (like observed clicks) to adjust payments and restore truthfulness. While our information structure differs (advertisers' state-dependent conversion rates are public information in our paper), our work addresses the same high-level challenge for the platform: designing a mechanism that must jointly elicit preferences and handle uncertainty about a common state.

2 Preliminaries

We consider the problem of a platform with a single ad slot for sale, to one of n strategic advertisers. We will refer to the ad slot as the *item*. The item is described by a state of nature $\theta \in \Theta$, for some finite set Θ. The state is initially unknown to both the platform and the advertisers, and it is sampled from a known distribution $\mathcal{D}$; we write $\mathcal{D}(\theta)$ for the probability that the state of nature is θ.

Each advertiser i has a private *base* value $v_i \geq 0$ for the item, which stays the same across all states θ. There is also a *state-dependent* conversion rate $\alpha_i(\theta) \geq 0$, known to the platform (as is the case in real-world systems), that represents the item's relevance for advertiser i given θ.[3] The value of advertiser i for an item with state θ is the base value multiplied by the conversion rate, i.e., $v_i \cdot \alpha_i(\theta)$. Advertisers have quasilinear utilities: when allocated an item with state θ for a payment p_i, advertiser i gets utility $u_i = v_i \cdot \alpha_i(\theta) - p_i$.

Before the item is allocated, the platform presents one of k possible questions. Presenting question ℓ reveals a random signal $\sigma \in \Sigma_\ell$ (for notational convenience, we assume wlog that Σ_ℓs are disjoint sets); σ is drawn from a conditional distribution over signals $Q_\ell(\cdot \mid \theta)$. The marginal signal distribution under question ℓ is $S_\ell(\sigma) := \sum_{\theta \in \Theta} Q_\ell(\sigma \mid \theta)\, \mathcal{D}(\theta), \sigma \in \Sigma_\ell$.

Upon observing σ, the posterior over states is $\mathcal{D}_{\ell,\sigma}(\theta) := \frac{\mathcal{D}(\theta)\, Q_\ell(\sigma|\theta)}{S_\ell(\sigma)}, \theta \in \Theta, S_\ell(\sigma) > 0$. Therefore, upon observing signal σ from question ℓ, the posterior expected conversion rate for advertiser i is $\alpha_i(\ell, \sigma) := \sum_{\theta \in \Theta} \alpha_i(\theta)\, \mathcal{D}_{\ell,\sigma}(\theta)$. We call $v_i\, \alpha_i(\ell, \sigma)$ the *effective value* of advertiser i given (ℓ, σ). We provide a continuation of Example 1.1 in Appendix A.

After the question is presented, and the signal is randomly drawn and revealed to all parties, the platform can allocate the item.

Overall, an instance of our problem is parameterized by: (1) the public prior distribution $\mathcal{D}$ over states, (2) public conditional distributions $Q_\ell(\sigma \mid \theta)$ for every question ℓ, signal σ, and state θ, (3) private base values v_i for each advertiser i, and (4) public conversion rates $\alpha_i(\theta)$ for each advertiser i and state θ. Given an instance, the timing of events is as follows: (1) Nature samples the item's state θ from $\mathcal{D}$. θ is not revealed to the platform or the advertisers. (2) The advertisers possibly interact with the platform. (3) The platform presents a question ℓ. (4) Nature draws a signal σ from $Q_\ell(\cdot \mid \theta)$, and reveals it to the platform and the advertisers. (5) The advertisers possibly interact with the platform. (6) The platform allocates the item to one of the advertisers (possibly in a randomized way), and charges payments.

Given this timing, the platform picks the rules of the game, and commits to them ex ante. The rules may condition on publicly observed signals σ and any messages solicited from advertisers in steps (2) and (5). In particular, the platform specifies (i) what input (if any) advertisers provide in steps (2) and (5); (ii) how the question ℓ is selected in step (3); and (iii) the allocation and payment rules used in step (6). The platform's goal is to maximize expected welfare: the expected sum, over the platform's (possibly randomized) choice of question and the induced signal, of each advertiser's base value times their conversion rate, weighted by the allocation probability. The goal of each advertiser is to maximize expected utility: its base value times conversion rate minus its payment.

In this paper, we study two different approaches for solving the platform's problem. First, in the *direct revelation* approach, the platform asks each advertiser to report their base values in step (2) (no interaction occurs in step (5)) and commits to a truthful mechanism. We define the setting in detail, and study such mechanisms, in Section 3. Second, we consider *modular two-stage mechanisms*: the platform commits to two auctions: one for picking the question in step (3), and one for allocating the item in step (6). Messages/bids are therefore solicited in both steps (2) and (5). Here, performance is evaluated at an equilibrium. We define the setting in detail, and study such mechanisms in Section 4.

3 Direct Revelation Mechanisms

In this section, we study the design of a single, end-to-end mechanism for both information acquisition and allocation. This corresponds to the direct revelation approach, where the platform commits to a set of rules, asks advertisers to report their private information (their base values), and then manages the entire process (which question to ask, who gets the item, how much to charge). The timing

[3] This can be, e.g., a click-through rate, or a conversion value.

of a direct mechanism is as follows: (1) Nature samples the item's state θ from $\mathcal{D}$. θ is not revealed to the platform or the advertisers. (2) Each advertiser i submits to the platform a bid b_i (possibly different than their base value v_i). (3) The platform presents a question ℓ. (4) Nature draws a signal σ from $Q_\ell(. \mid \theta)$, and reveals it to the platform and the advertisers. (5) The platform allocates the item to one of the advertisers (possibly in a randomized way), and charges payments.

A direct mechanism $\mathcal{M} = (\lambda, q, p)$ consists of three functions, that depend on the reported bids $b = (b_1, b_2, \ldots, b_n)$, from the single interaction with the advertisers in step (2) above:

- The *suggestion rule* $\lambda(b)$ is a probability distribution over the available questions; we write $\lambda_\ell(b)$ for the probability that the mechanism picks question ℓ. It holds that $\sum_{\ell=1}^{k} \lambda_\ell(b) = 1$, for all b.
- The *allocation rule* $q(b, \sigma)$: For a given report b and a realized signal σ (drawn from $Q_\ell(\cdot \mid \theta)$, where ℓ is the question sampled from $\lambda(b)$), $q_i(b, \sigma)$ is the probability that advertiser i receives the item. It holds that $\sum_{i=1}^{n} q_i(b, \sigma) \leq 1$, for all b and σ.
- The *payment rule* $p(b, \sigma)$: For a given report b and a realized signal σ, $p_i(b, \sigma)$ is the payment charged to advertiser i.

For deterministic mechanisms, $\lambda_\ell(b)$/$q_i(b, \sigma)$ are indicators for ℓ being chosen/i winning the item.

Let $x_i(b) = E_{\ell \sim \lambda(b), \sigma \sim S_\ell(.)}[\alpha_i(\ell, \sigma) \cdot q_i(b, \sigma)]$ be the expected *delivered* conversion rate for advertiser i, where the expectation is over the randomness of the mechanism and the random signal. Intuitively, if $\alpha_i(\ell, \sigma)$ is a click-through rate, $x_i(b)$ is the "expected number of clicks" allocated to advertiser i. Similarly, slightly overloading notation, let $p_i(b) = E_{\ell \sim \lambda(b), \sigma \sim S_\ell(.)}[p_i(b, \sigma)]$ be the expected payment of advertiser i. The interim utility of advertiser i, with a true base value v_i, who submits a bid b_i, and others report b_{-i} is therefore $U_i(v_i \to b_i; b_{-i}) = v_i \cdot x_i(b_i; b_{-i}) - p_i(b_i; b_{-i})$. Our notion of truthfulness is *dominant strategy incentive compatibility*. A mechanism is dominant strategy incentive compatible (DSIC) if truthtelling is a dominant strategy: $U_i(v_i \to v_i; b_{-i}) \geq U_i(v_i \to b_i; b_{-i})$ for all v_i, b_i and b_{-i}. We are interested in designing a DSIC mechanism that maximizes expected social welfare.

3.1 Characterizing the Optimal End-to-End Auction

Our first result establishes that the welfare-maximizing direct mechanism is truthful. This is achieved by defining the suggestion and allocation rules to greedily maximize the reported welfare, and then applying the classic Vickrey-Clarke-Groves (VCG) payment structure.

Lemma 3.1. *The direct mechanism $\mathcal{M}^* = (\lambda^*, q^*, p^{VCG})$ that maximizes expected social welfare is Dominant Strategy Incentive Compatible (DSIC), where:*

- *The suggestion rule $\lambda^*(b)$ deterministically chooses the question $\ell^* = \ell^*(b)$ that maximizes the ex-ante expected welfare, i.e., $\ell^* = \text{argmax}_\ell \, E_{\sigma \sim S_\ell(.)}[max_{j=1,\ldots,n} b_j \cdot \alpha_j(\ell, \sigma)]$.*
- *The allocation rule $q^*(b, \sigma)$ allocates the item to the advertiser i^* with the highest* posterior *value after suggestion $\ell^*(b)$ is presented, and signal σ is observed: $i^*(b, \sigma) = \text{argmax}_{j \in \{1,\ldots,n\}} \, b_j \cdot \alpha_j(\ell^*(b), \sigma)$. That is, $q_{i^*} = 1$ and $q_j = 0$ for $j \neq i^*$.*
- *The payment rule p^{VCG} charges each advertiser i the externality they impose on all other advertisers: the expected welfare of the optimal solution without i in the instance, minus the expected welfare of everyone other than i in the optimal solution with i.*

In Appendix A we provide an example where mechanism $\mathcal{M}^*$ is applied to Example 1.1.

The next lemma gives a simplified way of calculating the VCG payments in our problem. Specifically, the VCG payment of advertiser i can be decomposed into (1) an externality from stage 1, plus (2) an expected second price payment.

Lemma 3.2. *Let $\text{SW}_{-i}(\ell) := E_{\sigma \sim S_\ell}\big[\max_{j \neq i} b_j \, \alpha_j(\ell, \sigma)\big]$. Let $\ell^* = \ell^*(b)$ be the question chosen by $\mathcal{M}^* = (\lambda^*, q^*, p^{VCG})$. Then*

$$p_i^{VCG}(b) = \underbrace{\big(\max_{\ell} \text{SW}_{-i}(\ell) - \text{SW}_{-i}(\ell^*)\big)}_{\textit{Stage–1 externality}} + \underbrace{E_{\sigma \sim S_{\ell^*}}\Big[\max_{j \neq i} b_j \, \alpha_j(\ell^*, \sigma) \, q_i^*(b, \sigma)\Big]}_{\textit{expected Stage–2 second price payment}}.$$

The proofs of Lemmas 3.1 and 3.2 are deferred to Appendix B.

4 Modular Mechanisms: Decoupling Questions and Allocation

In this section, we study *modular mechanisms* where the platform decouples the problem into two distinct stages. First, it runs an auction to decide which question to present. Second, after the signal from that question is publicly observed, it runs a separate auction to allocate the item. This approach is simpler to implement, but requires advertisers to act strategically.

The timing of a modular mechanism is as follows: (1) Nature samples the item's state θ from $\mathcal{D}$; θ is not revealed to the platform or the advertisers. (2) Each advertiser i submits a vector of bids $b_i^Q = (b_{i,1}^Q, \ldots, b_{i,k}^Q)$, where $b_{i,\ell}^Q$ is its bid for having question ℓ presented. (3) The platform selects a question ℓ and charges payments based on the bids $\{b_i^Q\}_{i=1}^n$. (4) Nature draws a signal σ from $Q_\ell(\cdot|\theta)$, and reveals it to the platform and the advertisers. (5) Each advertiser i submits a bid b_i^A for the item. (6) The platform allocates the item and charges payments based on the bids $\{b_i^A\}_{i=1}^n$.

Analytically, we will treat this as a **one-shot normal-form game**: before play begins, each advertiser commits to a contingent plan for both stages. That is, advertisers *simultaneously* submit (i) a bid for each question and (ii) a bidding function for the item, mapping (ℓ, σ) to a stage-2 bid. The subsequent "stages" are merely the platform executing those plans after the signal is realized. Later in this section, we prove various results about equilibria of this normal form game; we note that our results can be extended to hold for the corresponding subgame-perfect equilibria for the corresponding sequential game.

A modular mechanism $\mathcal{M} = (\mathcal{M}_Q, \mathcal{M}_A)$ is composed of two separate mechanisms. The stage 1 mechanism, $\mathcal{M}_Q$, takes as input bids $\{b_i^Q\}$, and consists of: (i) a *question selection rule* $\lambda(b^Q)$, which gives a probability distribution over the available questions; we write $\lambda_\ell(b^Q)$ for the probability that the mechanism picks question ℓ and have that $\sum_\ell \lambda_\ell(b^Q) = 1$. (ii) a stage 1 payment rule $p^Q(b^Q)$ that charges a payment $p_i^Q(b^Q)$ to each advertiser i. The stage 2 mechanism, $\mathcal{M}_A$, takes as input bids $\{b_i^A\}$, and consists of: (i) an *item selection rule* $q(b^A; \ell, \sigma)$, that allocates the item to advertiser i with probability $q_i(b^A; \ell, \sigma)$, and (ii) a stage 2 payment rule $p^A(b^A; \ell, \sigma)$ that charges each advertiser i a payment $p_i^A(b^A; \ell, \sigma)$.

4.1 Equilibrium Selection and Behavioral Assumptions

We evaluate these mechanisms under *rational behavior* assumptions. However, the specific equilibrium refinement we apply depends on the strategic structure of the mechanism being analyzed.

A strategy s_i for an advertiser i specifies what they will do at every interaction point. That is, a strategy s_i consists of (1) a stage 1 bidding vector $b_i^Q = (b_{i,1}^Q, \ldots, b_{i,k}^Q)$, which specifies their bid for each question ℓ, and (2) a bidding function $b_i^A(\ell, \sigma)$ which specifies what they will bid for the item for every possible question ℓ and signal σ that could be revealed. Let $s = (s_1, \ldots, s_n)$ be a strategy profile for the advertisers. This profile determines the bids in both stages, and thus the final outcome and advertisers' utilities. Let $U_i(s_i; s_{-i})$ be the expected utility for advertiser i when they play strategy s_i and all other advertisers play according to strategies in s_{-i}, where the expectation is over the strategies of others, the environment (random state of the item, signal), and randomness in the modular mechanism.

A strategy profile s^* is a Nash equilibrium if no advertiser has a profitable unilateral deviation; that is, if for every advertiser i, and any strategy s_i, $U_i(s_i^*; s_{-i}^*) \geq U_i(s_i; s_{-i}^*)$. The worst case inefficiency of a Nash equilibrium in a model is captured by a concept called the *Price of Anarchy* [KP09], the ratio of the social welfare of the optimal direct revelation outcome to that of the worst-case Nash equilibrium.

It is well-understood that, without any additional assumptions, Price of Anarchy bounds for certain auction environments can be unreasonably pessimistic. For example, the Price of Anarchy of a (one-shot) second price auction is unbounded: it is an equilibrium for an agent with a value of zero to bid infinity, while every other agent bids zero. Therefore, it is standard in this literature to assume that bidders are conservative, e.g., don't bid more than their value [LT10, LB10, BR11, CKKK11, DKMST13, CKS16]. To ensure our results are robust to such pathologies, we also restrict our analysis to *conservative* advertisers.

Definition 4.1 (Conservative advertisers)**.** *An advertiser i is* conservative *if they always bid at most their expected value for an outcome. Specifically,*

1. *In stage 2, an advertiser never bids more than their value. I.e., $b_i^A(\ell, \sigma) \leq v_i$, for all ℓ, σ.*
2. *In stage 1, an advertiser never bids on a question more than their expected value for the item in stage 2. I.e., $b_{i,\ell}^Q \leq v_i \cdot E_{\theta \sim \mathcal{D}}[\alpha_i(\theta)]$.*

For mechanisms that offer clear strategic guidance, such as those with dominant strategies, we can adopt a stricter assumption based on standard rationality (noting that restricting the space of available strategies only makes the negative results stronger). Overall, we will assume the following behavior for a *rational* advertiser:

- **Mechanisms with Dominant Strategies:** For mechanisms that admit a weakly dominant strategy (such as VCG), we will assume advertisers are sophisticated enough to play it. Formally, we analyze equilibria in *undominated strategies*. This assumption strengthens our negative results: if the PoA is unbounded under undominated strategies, it is unbounded for conservative bidders.
- **Mechanisms without Dominant Strategies:** For mechanisms where no dominant strategy exists (such as First-Price or All-Pay auctions), we will assume that advertisers are conservative (as defined in Definition 4.1). Note that, if the Price of Anarchy is unbounded for conservative bidders, then it is unbounded in general.

4.2 A Natural Modular Mechanism: VCG per Stage

We instantiate the general modular framework with a specific, natural choice for the two mechanisms. Specifically, we suggest running VCG for each stage separately.

Working backwards, in the second stage VCG boils down to a second price auction on "effective" values. Concretely, after a question ℓ has been chosen (in step (3)) and a signal σ has been publicly observed (in step (4)), the platform holds the following auction: allocate the item to the advertiser with the highest *effective* bid, i.e. the advertiser i^* with the highest $b_i^A \cdot \alpha_i(\ell, \sigma)$, and charge them the second-highest effective bid $p_{i^*}^A = \max_{j \neq i^*} b_j^A \cdot \alpha_j(\ell, \sigma)$.

Anticipating the truthful and efficient outcome of the second stage auction, each advertiser can calculate its expected utility for any given question ℓ. Concretely, for a fixed question ℓ, an advertiser can simulate a draw σ from $S_\ell(.)$, and the resulting payoff from the subsequent second price auction. This induces a well-defined expected utility $R_i(\ell; v)$, for each advertiser i and question ℓ. Therefore, the stage 1 problem boils down to picking a mechanism for the following "public projects" problem: there are k projects to choose from (the questions), and n agents (the advertisers), each with a private value $R_i(\ell; v_i)$ for each project. For this problem, VCG is again the natural, truthful solution: pick the project (question) that maximizes social welfare, and charge each agent its externality. Concretely, given a bid $b_{i,\ell}^Q$ from each advertiser i for each question ℓ: present the question ℓ^* with maximum sum of bids, i.e. $\ell^* = \text{argmax}_\ell \sum_{i=1}^n b_{i,\ell}^Q$, and charge each advertiser i its externality $p_i^Q = \left(\max_\ell \sum_{j \neq i} b_{j,\ell}^Q\right) - \left(\sum_{j \neq i} b_{j,\ell^*}^Q\right)$.

In Appendix A we provide an example of this mechanism, applied to the instance from Example 1.1.

4.3 Unbounded Price of Anarchy in Modular Mechanisms

We first prove that the mechanism from Section 4.2, running VCG in every stage, has a natural undominated pure Nash equilibrium: every advertiser bids their (true) expected utility for each question, as well as their true base value. We defer the proof of Theorem 4.1 to Appendix B.

Theorem 4.1 (Existence of an undominated pure Nash equilibrium)**.** *Consider the modular mechanism $\mathcal{M} = (\mathcal{M}_Q^{VCG}, \mathcal{M}_A^{VCG})$ where $\mathcal{M}_A^{VCG}$ is the (per-signal) VCG auction on effective values and $\mathcal{M}_Q^{VCG}$ is VCG for choosing the question, as defined in Section 4.2. Define, for each advertiser i and question ℓ,*

$$R_i(\ell; v) \ := \ E_{\sigma \sim S_\ell}\Big[\, u_i^A\big(v; \ell, \sigma\big)\,\Big],$$

where $u_i^A(v;\ell,\sigma)$ is i's (quasilinear) utility in $\mathcal{M}_A^{VCG}$ when the realized question is ℓ, the realized signal is σ, and everyone bids their base value in stage 2. Then the strategy profile

$$b_i^A(\ell,\sigma) = v_i \quad \textit{for all } i,\ell,\sigma, \qquad b_{i,\ell}^Q = R_i(\ell;v) \quad \textit{for all } i,\ \ell$$

is an undominated pure Nash equilibrium.

The equilibrium outcome in Theorem 4.1 might differ from the welfare-maximizing outcome (the outcome of the mechanism as described in Section 3). Specifically, while the stage 2 outcome of the equilibrium of Theorem 4.1 maximizes welfare for every realized question ℓ and signal σ, the stage 1 question selected maximizes expected *quasilinear utility*

$$\ell_{eq}^* = \operatorname*{argmax}_\ell \sum_{i=1}^n R_i(\ell;v) = \operatorname*{argmax}_\ell \sum_{i=1}^n E_{\sigma\sim S_\ell}\Big[\, u_i^A\big(v;\ell,\sigma\big)\,\Big].$$

However, the welfare-maximizing outcome selects the question that maximizes expected *value*

$$\ell_{opt}^* = \operatorname*{argmax}_\ell E_{\sigma\sim S_\ell(.)}[\max_{j=1,\ldots,n} v_j\,\alpha_j(\ell,\sigma)].$$

Therefore, it is natural to ask how inefficient this equilibrium outcome can be. And, more generally, how bad is the Price of Anarchy of the proposed mechanism $\mathcal{M} = (\mathcal{M}_Q^{VCG}, \mathcal{M}_A^{VCG})$?

Our next theorem shows that not only do inefficient equilibria exist, but even the pure Nash equilibrium suggested in Theorem 4.1 has infinite Price of Anarchy.

Theorem 4.2. *The Price of Anarchy of the mechanism $\mathcal{M} = (\mathcal{M}_Q^{VCG}, \mathcal{M}_A^{VCG})$ in undominated strategies is unbounded. Moreover, for all $c \geq 1$, there exists an instance such that the optimal welfare is at least c times the welfare of the equilibrium from Theorem 4.1.*

Proof. Fix an integer $m \geq 3$ and a $\delta > 0$ (fixed later in this proof). Our instance is as follows:

- States and questions:
  1. $\Theta = \{1, 2, \ldots, m\}$ with the uniform prior $\mathcal{D}(\theta = t) = 1/m$.
  2. An informative question $\ell = 1$: $\Sigma_1 = \{\sigma_1, \ldots, \sigma_m\}$ with $Q_1(\sigma_t \mid \theta = t) = 1$. Thus, $\mathcal{D}_{1,\sigma_t}$ is a point mass on state t.
  3. An uninformative question $\ell = 2$: $\Sigma_2 = \{\bar{\sigma}\}$ with $Q_2(\bar{\sigma} \mid \theta) = 1$ for all θ. Thus, $\mathcal{D}_{2,\bar{\sigma}} = \mathcal{D}$.
- Advertisers:
  1. There are $n = m$ advertisers, each with a base value $v_i = 1$.
  2. For each state $i \in \{1, \ldots, n\}$, $\alpha_i(i) = 1$. That is, advertiser i is the "primary" advertiser for state i.
  3. For every state $i \in \{1, \ldots, n\}$, $i \neq 3$, $\alpha_{i-1}(i) = 1 - \delta$ (where $i - 1 = n$ for $i = 1$). That is, advertiser i is the "secondary" advertiser for state $i + 1$.
  4. For state $i = 3$, $\alpha_1(3) = 1 - \delta$ and $\alpha_2(3) = 0$. That is, advertiser 1 is the "secondary" advertiser for state 3 as well (also the secondary for state 2, and the primary for state 1), while advertiser 2 is not the secondary advertiser for any state.
  5. All remaining conversion rates are 0: $\alpha_i(t) = 0$ for all other pairs (i, t).

This construction ensures that we have exactly two advertisers with positive bids in each state (which creates competition under the revealing question). We now compute induced (signal) conversion rates and compare the two questions.

Induced conversion rates. Under question $\ell = 1$, for every realized σ_t the posterior is a point mass at state t, therefore, $\alpha_i(1, \sigma_t) = \alpha_i(t)$ for every state t (and signal σ_t) and advertiser i. On the other hand, under the uninformative question $\ell = 2$, the posterior equals the prior; therefore,

$\alpha_i(2,\bar{\sigma}) = E_{\theta\sim\mathcal{D}}\big[\alpha_i(\theta)\big] = \frac{1}{m}\sum_{t=1}^{m}\alpha_i(t)$. From our construction, we have:

$$\alpha_1(2,\bar{\sigma}) = \frac{1+2(1-\delta)}{m} = \frac{3-2\delta}{m}$$
$$\alpha_2(2,\bar{\sigma}) = \frac{1}{m}$$
$$\alpha_i(2,\bar{\sigma}) = \frac{1+(1-\delta)}{m} = \frac{2-\delta}{m} \quad \text{for all } i \notin \{1,2\}$$

Stage 2 outcome. Given any realized (ℓ,σ), $\mathcal{M}_A^{VCG}$ is a single-item VCG auction on effective values $b_i^A\,\alpha_i(\ell,\sigma) = v_i\,\alpha_i(\ell,\sigma) = \alpha_i(\ell,\sigma)$ (truthful bidding is a dominant strategy, and $v_i = 1$).

If $\ell = 1$, in every state t, there are two positive effective values: 1 from the primary advertiser (advertiser t) and $1-\delta$ from the secondary advertiser (advertiser $t-1$ for $t \notin \{1,3\}$, advertiser 1 for $t = 3$, and advertiser n for $t = 1$). Thus, for any state t, we have that the welfare (the value generated) is 1, but the utility of the winning advertiser (the primary advertiser) is δ. Taking an expectation over the uniform prior we have that $R_i(1;v) = \delta/m$ for all i, where $R_i(\ell;v) = E_{\sigma\sim S_\ell}[u_i^A(v;\ell,\sigma)]$ is advertiser i's utility under question ℓ.

If $\ell = 2$, there is a single signal $\bar{\sigma}$ with effective values $\alpha_1(2,\bar{\sigma}) = \frac{3-2\delta}{m}$, $\alpha_2(2,\bar{\sigma}) = \frac{1}{m}$, and $\alpha_i(2,\bar{\sigma}) = \frac{2-\delta}{m}$ for all $i \notin \{1,2\}$. Therefore, under VCG (a second price auction) with these bids, we have that the (expected) value is $\frac{3-2\delta}{m}$ (for $\delta < 1$, we have $\frac{3-2\delta}{m} > \frac{2-\delta}{m}$), revenue $\frac{2-\delta}{m}$, and $R_i(2;v) = 0$ for $i \neq 1$, but $R_1(2;v) = \frac{3-2\delta}{m} - \frac{2-\delta}{m} = \frac{1-\delta}{m}$.

Stage 1 outcome. Recall that the equilibrium from Theorem 4.1 (which is an equilibrium in undominated strategies) prescribes that $b_{i,\ell}^Q = R_i(\ell;v)$, for all advertisers i and questions ℓ. Given the stage 2 calculations, we therefore have that the total bid on question $\ell = 1$ is $\sum_{i=1}^n R_i(1;v) = \delta$. On the other hand, for question $\ell = 2$ we have that the total bid is $\sum_{i=1}^n R_i(2;v) = R_1(2;v) = \frac{1-\delta}{m}$.

By picking δ such that $\frac{1-\delta}{m} > \delta$ (e.g., $\delta = 1/m^2$), we have that the mechanism picks question $\ell = 2$. Therefore, in the prescribed equilibrium for the mechanism $\mathcal{M} = (\mathcal{M}_Q^{VCG}, \mathcal{M}_A^{VCG})$, the overall outcome is to pick question $\ell = 2$, which results in advertiser 1 always winning in stage 2, for a total (expected) welfare of $\frac{3-2\delta}{m}$. On the other hand, presenting question $\ell = 1$ results in the "primary" advertiser winning the item in every state t, for a total (expected) welfare of 1. □

A natural attempt to bypass the negative result of Theorem 4.2 is to keep the stage 2 mechanism fixed to a second price auction (which is a synonym to welfare maximization), and make tweaks to the stage 1 mechanism. For example, we could keep the same question selection rule (select the question with the highest total bid) and change the payment rule to first-price or all-pay payments. As we show next, the result of Theorem 4.2 seems robust to such changes. The proof of Theorem 4.3 is deferred to Appendix B.

Theorem 4.3 (Robustness of Unbounded PoA)**.** *Even for conservative bidders (Definition 4.1), the Price of Anarchy remains unbounded if the stage 1 VCG auction is replaced by a mechanism that selects the question with the highest sum of bids, and uses one of the following payment rules: (i)* ***First-Price:*** *The advertiser who submitted the single highest bid for the winning question pays that bid (and no other advertiser pays), or (ii)* ***All-Pay:*** *Every advertiser pays their submitted bid for the winning question.*

4.4 Modular Mechanisms with Better Equilibria

The negative results in Theorems 4.2 and 4.3 are crucially using the fact that, under a second price auction, the welfare can be high but the total advertiser utility might be zero (e.g., when the highest bid is equal to the second highest bid). When this occurs in the second stage, the advertisers are not incentivized to bid high on the "right" questions in stage 1, and therefore these questions are not selected. To avoid such situations, we need second stage mechanisms that give advertisers high utility even when the competition is high. The extra advertiser utility can only come from the platform's revenue. A line of work in mechanism design [Cav06, GC09, Guo12, GN11, Guo11] studies the problem of re-allocating the revenue generated back to the participating agents (this is motivated by, for instance, situations where agents are trying to re-distribute resources among

Table 1: Summary of Results for Modular and Unified Mechanisms

Mechanism	Stage 1 Format	Assumption	PoA	
VCG-per-stage	VCG	Undominated Nash	Unbounded	Theorem 4.2
Modular FP	First-Price	Conservative Bidders	Unbounded	Theorem 4.3
Modular AP	All-Pay	Conservative Bidders	Unbounded	Theorem 4.3
Redistribution	VCG	Undominated Nash	$1 + \frac{2}{n-2}$ (for $n \geq 3$)	Lemma 4.1

themselves). These are so-called *mechanisms with redistribution*. In the remainder of this section we describe and analyze a modular mechanism for our problem that replaces the second stage VCG auction with the so-called Bailey-Cavallo mechanism [Bai97, Cav06].

Stage 2: Ad Allocation (Bailey-Cavallo mechanism). Working backwards, after a question ℓ has been chosen and a signal σ has been publicly observed, the platform runs an auction with redistribution, as proposed by Bailey [Bai97] and Cavallo [Cav06]. First, the platform determines the allocation and VCG payments based on effective bids. The item is allocated to the advertiser i^* with the highest effective bid, i.e., $i^* = \operatorname{argmax}_{i \in \{1,\dots,n\}} b_i^A \cdot \alpha_i(\ell, \sigma)$. The base payment is the standard VCG payment (the second-highest effective bid): $p_{i^*}^{\text{VCG}} = \max_{i \neq i^*} b_i^A \cdot \alpha_i(\ell, \sigma)$.

Crucially, the mechanism then applies a *redistribution rebate* to all advertisers. Each advertiser i receives a rebate equal to an equal share of the VCG revenue that would have been generated if i were not present. Let R_{-i} be the VCG revenue of the auction instance where advertiser i is removed. For i^* and the agent with the second highest effective bid, R_{-i} is equal to the third highest effective bid; for everyone else, R_{-i} is equal to the second highest effective bid. The final payment for advertiser i is: $p_i^A = p_i^{\text{VCG}} - \frac{R_{-i}}{n}$. Note that for non-winning bidders (i.e., $i \neq i^*$), $p_i^{\text{VCG}} = 0$, so they simply receive a payment of $\frac{R_{-i}}{n}$.

Stage 1: Question Selection (Standard VCG) The first stage proceeds identically to the modular VCG mechanism described in Section 4.2. The platform solicits bids $b_{i,\ell}^Q$ from each advertiser i for each question ℓ. It then selects the question ℓ^* that maximizes the sum of reported bids: $\ell^* = \operatorname{argmax}_\ell \sum_{i=1}^n b_{i,\ell}^Q$. Finally, the platform charges each advertiser i their standard VCG externality for the chosen question, exactly as defined in Section 4.2.

The proof of the following lemma is deferred to Appendix B.

Lemma 4.1. *For $n \geq 3$, the Price of Anarchy in undominated strategies of the modular mechanism that runs Bailey-Cavallo mechanism in stage 2 and VCG in stage 1, is at most $1 + \frac{2}{n-2}$.*

The results for the various modular designs are summarized in Table 1.

While Lemma 4.1 provides a positive theoretical result, showing that it is possible to achieve a constant Price of Anarchy, we do not necessarily advocate for the adoption of the VCG-plus-redistribution modular mechanism in practice.

First, the mechanism suffers from various practical hurdles: it is complex to explain to advertisers, and participating with zero bids is, implicitly, incentivized (since zero-bidding advertisers receive strictly positive utility). Second, the revenue for the platform is strictly lower compared to pure VCG; the sacrifice in net revenue might pose an additional practical hurdle for real-world adoption. Third, and perhaps more subtly, the strength of this positive result depends heavily on the assumption that agents play the pair of dominant strategies. For stage 2, this assumption is easier to justify, but in stage 1 computing the "right" bid is somewhat more complex (and therefore harder to justify in practice). And, even under the weaker behavioral assumption of conservative bidding (and despite the fact that this assumption seems somewhat necessary for analyzing standard auctions), the distinction between robust mechanism design and trivial solutions blurs. For instance, in a standard one-shot single-item auction, a naive mechanism that allocates the item to the highest bidder but charges zero payments would also technically achieve a Price of Anarchy of 1 under a no-overbidding assumption (like conservative bidding).

The value of Lemma 4.1 is therefore not that it offers a perfect engineering solution, but that it demonstrates that positive results for modular mechanisms are possible, provided one is willing to accept some additional assumptions.

References

[Bai97] Martin J Bailey. The demand revealing process: to distribute the surplus. *Public Choice*, 91(2):107–126, 1997.

[BBD+25] Dirk Bergemann, Marek Bojko, Paul Duetting, Renato Paes Leme, Haifeng Xu, and Song Zuo. Data-driven mechanism design: Jointly eliciting preferences and information. In *Proceedings of the 26th ACM Conference on Economics and Computation*, EC '25, page 507, New York, NY, USA, 2025. Association for Computing Machinery.

[BMP25] Martino Banchio, Aranyak Mehta, and Andres Perlroth. Ads in conversations. In *Proceedings of the 26th ACM Conference on Economics and Computation*, pages 350–350, 2025.

[BR11] Kshipra Bhawalkar and Tim Roughgarden. Welfare guarantees for combinatorial auctions with item bidding. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 700–709. SIAM, 2011.

[Cav06] Ruggiero Cavallo. Optimal decision-making with minimal waste: Strategyproof redistribution of vcg payments. In *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, pages 882–889, 2006.

[CKKK11] Ioannis Caragiannis, Christos Kaklamanis, Panagiotis Kanellopoulos, and Maria Kyropoulou. On the efficiency of equilibria in generalized second price auctions. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 81–90, 2011.

[CKS16] George Christodoulou, Annamária Kovács, and Michael Schapira. Bayesian combinatorial auctions. *Journal of the ACM (JACM)*, 63(2):1–19, 2016.

[DFK+24] Avinava Dubey, Zhe Feng, Rahul Kidambi, Aranyak Mehta, and Di Wang. Auctions with llm summaries. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '24, page 713–722, New York, NY, USA, 2024. Association for Computing Machinery.

[DKMST13] Bart De Keijzer, Evangelos Markakis, Guido Schäfer, and Orestis Telelis. Inefficiency of standard multi-unit auctions. In *European Symposium on Algorithms*, pages 385–396. Springer, 2013.

[DMPL+24] Paul Duetting, Vahab Mirrokni, Renato Paes Leme, Haifeng Xu, and Song Zuo. Mechanism design for large language models. In *Proceedings of the ACM Web Conference 2024*, pages 144–155, 2024.

[FHRS25] Soheil Feizi, MohammadTaghi Hajiaghayi, Keivan Rezaei, and Suho Shin. Online advertisements with llms: Opportunities and challenges. *ACM SIGecom Exchanges*, 22(2):66–81, 2025.

[GC09] Mingyu Guo and Vincent Conitzer. Worst-case optimal redistribution of vcg payments in multi-unit auctions. *Games and Economic Behavior*, 67(1):69–98, 2009.

[GN11] Sujit P Gujar and Yadati Narahari. Redistribution mechanisms for assignment of heterogeneous objects. *Journal of Artificial Intelligence Research*, 41:131–154, 2011.

[Gooa] Google. Adsense for search (afs) policies.

[Goob] Google. Our approach - how google search works.

[Goo24] Google Search Central. New search experiences in eea: Rich results, aggregator units, and refinement chips. `https://developers.google.com/search/blog/2024/02/search-experiences-in-eea`, February 2024. Accessed: 2025-10-27.

[Goo25] Google. Related search for your content pages - google adsense help, 2025.

[Guo11] Mingyu Guo. Vcg redistribution with gross substitutes. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 25, pages 675–680, 2011.

[Guo12] Mingyu Guo. Worst-case optimal redistribution of vcg payments in heterogeneous-item auctions with unit demand. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 745–752, 2012.

[HLRS24] MohammadTaghi Hajiaghayi, Sébastien Lahaie, Keivan Rezaei, and Suho Shin. Ad auctions for llms via retrieval augmented generation. *Advances in Neural Information Processing Systems*, 37:18445–18480, 2024.

[KP09] Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. *Computer Science Review*, 3(2):65–69, 2009.

[LB10] Brendan Lucier and Allan Borodin. Price of Anarchy for greedy auctions. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, pages 537–553. SIAM, 2010.

[LT10] Renato Paes Leme and Eva Tardos. Pure and bayes-nash Price of Anarchy for generalized second price auction. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 735–744. IEEE, 2010.

[Mic25] Microsoft. How bing delivers search results, March 2025.

[MTK24] Tommy Mordo, Moshe Tennenholtz, and Oren Kurland. Sponsored question answering. In *Proceedings of the 2024 ACM SIGIR International Conference on Theory of Information Retrieval*, ICTIR '24, page 167–173, New York, NY, USA, 2024. Association for Computing Machinery.

[Per24] Perplexity AI. Why we're experimenting with advertising. `https://www.perplexity.ai/hub/blog/why-we-re-experimenting-with-advertising`, 2024. Accessed: September 28, 2025.

[SCS25] Ermis Soumalias, Michael J. Curry, and Sven Seuken. Truthful aggregation of llms with an application to online advertising, 2025.

A Additional Examples

Example A.1 (Running shoes (continuation of Example 1.1))**.** *Recall that $\Theta = \{(\theta_{terr}, \theta_{exp}) \in \{0,1\}^2\}$ with uniform prior $\mathcal{D}$. We have two advertisers with private base values $v_1 = 50$, $v_2 = 30$, and public conversion rates:*

$$\alpha_1(\theta) = \begin{cases} 0.9 & \text{if } (\theta_{terr}, \theta_{exp}) = (1,1), \\ 0.3 & \text{if } (\theta_{terr}, \theta_{exp}) = (1,0), \\ 0 & \text{if } \theta_{terr} = 0, \end{cases} \qquad \text{and} \qquad \alpha_2(\theta) = \begin{cases} 0.8 & \text{if } (\theta_{terr}, \theta_{exp}) = (0,0), \\ 0.4 & \text{if } (\theta_{terr}, \theta_{exp}) = (0,1), \\ 0 & \text{if } \theta_{terr} = 1. \end{cases}$$

There are two questions:

1. ***Terrain question**: ℓ_{terr} ("I am looking for trail running shoes?"), that has signals $\Sigma_{terr} = \{click, no\text{-}click\}$ and*

$$Q_{terr}(click \mid \theta) = \mathbf{1}\{\theta_{terr} = 1\} \qquad Q_{terr}(no\text{-}click \mid \theta) = \mathbf{1}\{\theta_{terr} = 0\}.$$

   Thus, $S_{terr}(click) = S_{terr}(no\text{-}click) = \frac{1}{2}$. The posteriors are: $\mathcal{D}_{terr,click}$ is uniform over $\{(1,1),(1,0)\}$, and $\mathcal{D}_{terr,no\text{-}click}$ is uniform over $\{(0,1),(0,0)\}$. The posterior expected conversion rates are

$$\alpha_1(terr, click) = \tfrac{0.9+0.3}{2} = 0.6, \qquad \alpha_1(terr, no\text{-}click) = 0,$$
$$\alpha_2(terr, click) = 0, \qquad \alpha_2(terr, no\text{-}click) = \tfrac{0.4+0.8}{2} = 0.6.$$

   Therefore, the effective value of advertiser 1 is $50 \cdot 0.6 = 30$ when the signal is "click," and 0 otherwise. The effective value of advertiser 2 is 0 when the signal is "click," and $30 \cdot 0.6 = 18$ otherwise. Therefore, ex-ante, ℓ_{terr} induces a lottery that with probability $1/2$ gives effective values 30 and 0, and with probability $1/2$ the effective values are 0 and 18.

2. ***Targeted question**: ℓ_{tgt} ("I am an experienced trail runner"), with $\Sigma_{tgt} = \{click, no\text{-}click\}$ and*

$$Q_{tgt}(click \mid \theta) = \mathbf{1}\{\theta = (1,1)\} \qquad Q_{tgt}(no\text{-}click \mid \theta) = 1 - Q_{tgt}(click \mid \theta).$$

   Thus, $S_{tgt}(click) = \frac{1}{4}$ and $S_{tgt}(no\text{-}click) = \frac{3}{4}$. The posteriors are: $\mathcal{D}_{tgt,click}$ is a point mass on $(1,1)$, and $\mathcal{D}_{tgt,no\text{-}click}$ is uniform over $\{(1,0),(0,1),(0,0)\}$. The posterior expected conversion rates are

$$\alpha_1(tgt, click) = 0.9, \quad \alpha_1(tgt, no\text{-}click) = \tfrac{0.3+0+0}{3} = 0.1,$$
$$\alpha_2(tgt, click) = 0, \quad \alpha_2(tgt, no\text{-}click) = \tfrac{0+0.4+0.8}{3} = \tfrac{1.2}{3} = 0.4.$$

   Therefore, the effective value of advertiser 1 is $50 \cdot 0.9 = 45$ when the signal is "click," and $50 \cdot 0.1 = 5$ otherwise. The effective value of advertiser 2 is 0 when the signal is "click," and $30 \cdot 0.4 = 12$ otherwise. Therefore, ex-ante, ℓ_{tgt} induces a lottery that with probability $1/4$ gives effective values 45 and 0, and with probability $3/4$ the effective values are 5 and 12.

Example A.2 (Mechanism $\mathcal{M}^*$ applied to Example A.1)**.** *Consider applying $\mathcal{M}^*$ to the instance from Example A.1 given (truthful) bids $b_1 = 50$, $b_2 = 30$. Given these bids, the mechanism first chooses a question to present.*

***Choosing the question** $\ell^*(b)$. As discussed in Example A.1, question ℓ_{terr} induces a lottery that, with probability $1/2$, gives effective values (base value times posterior expected conversion rate) 30 and 0, and with probability $1/2$, the effective values are 0 and 18. The expected welfare under ℓ_{terr} is $\frac{1}{2} \cdot 30 + \frac{1}{2} \cdot 18 = 24$. On the other hand, ℓ_{tgt} induces a lottery that, with probability $1/4$, gives effective values 45 and 0, and with probability $3/4$, the effective values are 5 and 12. Therefore, the expected welfare under ℓ_{tgt} is $\frac{1}{4} \cdot 45 + \frac{3}{4} \cdot 12 = 20.25$. Hence the mechanism picks $\ell^*(b) = \ell_{\text{terr}}$.*

***Allocating the item.** Since ℓ_{terr} is picked, then with probability $1/2$ the signal is "click," the induced effective values are 30 and 0, and advertiser 1 receives the item. Otherwise, the signal is "no-click," the induced values are 0 and 18, and advertiser 2 receives the item.*

***Payments.** The expected value of advertiser 1 in the current solution is 15: they receive the item when the signal from asking ℓ_{terr} is "click" (this is a probability $1/2$ event), and their effective value is 30. For advertiser 2, their expected value is 9: they receive the item when the signal from asking ℓ_{terr} is "no-click" (this is a probability $1/2$ event), and their effective value is 18 when this happens.*

What would happen if advertiser 1 did not participate? In that case, the question would be picked to optimize just for advertiser 2, who is indifferent between the two questions, since its expected value is (still) 9 *under both choices. This means that advertiser 1 does not impose any externality, so its payment is zero.*

What would happen if advertiser 2 did not participate? Similarly, the question would be picked to optimize just for advertiser 1, who is also indifferent between the two questions: ℓ_{terr} *gives an expected value of* $30/2 = 15$*, and* ℓ_{tgt} *gives an expected value of* $45/4 + 5 \cdot 3/4 = 15$*. Therefore that advertiser 2 does not impose any externality, so its payment is also zero.*

Example A.3. *Consider the instance from Example A.1.*

Assume that advertiser 1 submits: Stage 1 bids $b^Q_{1,terr} = 21$*, and* $b^Q_{1,tgt} = 20$*; Stage 2 bidding function* $b^A_1(\ell, \sigma) = v_1 = 50$ *for all* (ℓ, σ)*. Assume that advertiser 2 submits: Stage 1 bids* $b^Q_{2,terr} = 9$*, and* $b^Q_{2,tgt} = 12$*; Stage 2 bidding function* $b^A_2(\ell, \sigma) = v_2 = 30$ *for all* (ℓ, σ)*.*

Then the mechanism proceeds as follows. First, since $\sum_i b^Q_{i,tgt} = 20 + 12 > 21 + 9 = \sum_i b^Q_{i,terr}$*, the question chosen in the first stage is* $\ell^* = \ell_{tgt}$*. To compute the stage 1 VCG externalities for advertiser* 1 *we have:* $\max_\ell b^Q_{2,\ell} = \max\{9, 12\} = 12$ *and* $b^Q_{2,\ell^*} = 12$*, so* $p^Q_1 = 0$*. To compute the stage 1 VCG externalities for advertiser* 2 *we have:* $\max_\ell b^Q_{1,\ell} = \max\{21, 20\} = 21$ *and* $b^Q_{1,\ell^*} = 20$*, so* $p^Q_2 = 1$*. Then, a signal is realized in stage 2. Since* ℓ_{tgt} *was chosen:*

- *If* $\sigma =$ *click (prob.* $1/4$*), the effective bids are* $(b^A_1\alpha_1, b^A_2\alpha_2) = (50 \cdot 0.9,\ 30 \cdot 0) = (45, 0)$*. Advertiser 1 wins. The second-highest effective bid is* 0*, so the Stage 2 payment is* $p^A_1 = 0$*.*
- *If* $\sigma =$ *no-click (prob.* $3/4$*), the effective bids are* $(50 \cdot 0.1,\ 30 \cdot 0.4) = (5, 12)$*. Advertiser 2 wins and pays the second-highest effective bid* 5*, so* $p^A_2 = 5$*.*

B Missing Proofs

Proof of Lemma 3.1. Fix bids b. For any question ℓ and signal σ, the reported welfare from allocating to i is $b_i\,\alpha_i(\ell, \sigma)$. Hence, the maximum reported welfare for fixed ℓ is $E_{\sigma\sim S_\ell}\Big[\max_j b_j\,\alpha_j(\ell, \sigma)\Big]$, attained by the allocation rule $q^*(b, \sigma)$, by allocating, for each realized σ, to a maximizer of $b_j\,\alpha_j(\ell, \sigma)$. Optimizing over ℓ, gives the suggestion rule $\lambda^*(b)$ above; therefore, our rules maximize expected reported welfare over all feasible outcomes (choice of ℓ and signal-contingent allocation). Define

$$W_{-i}(b_{-i}) \;:=\; \max_{\ell\in[k]} E_{\sigma\sim S_\ell}\Big[\max_{j\neq i} b_j\,\alpha_j(\ell, \sigma)\Big].$$

Advertiser i's payment is then:

$$p^{VCG}_i(b) \;=\; W_{-i}(b_{-i}) - E_{\sigma\sim S_{\ell^*(b)}}\Big[\sum_{j\neq i} b_j\,\alpha_j\big(\ell^*(b), \sigma\big)\,q^*_j(b, \sigma)\Big].$$

Therefore, i's expected utility when its true base value is v_i and it reports b_i is: $U_i(v_i \to b_i; b_{-i}) = E_{\sigma\sim S_{\ell^*(b)}}\Big[v_i\,\alpha_i\big(\ell^*(b), \sigma\big)\,q^*_i(b, \sigma) + \sum_{j\neq i} b_j\,\alpha_j\big(\ell^*(b), \sigma\big)\,q^*_j(b, \sigma)\Big] \;-\; W_{-i}(b_{-i})$. Since (λ^*, q^*) selects, for each b, the outcome that maximizes $E_{\sigma\sim S_{\ell^*(b)}}[\sum_j b_j\alpha_j\big(\ell^*(b), \sigma\big)]$, the first term in $U_i(v_i \to b_i; b_{-i})$ is maximized when $b_i = v_i$. The second term ($W_{-i}(b_{-i})$) does not depend on i's report; therefore, $b_i = v_i$ is a dominant strategy. □

Proof of Lemma 3.2. Starting from the definition of $p^{VCG}_i(b)$ we have

$$p^{VCG}_i(b) = \max_\ell \mathrm{SW}_{-i}(\ell) \;-\; E_{\sigma\sim S_{\ell^*}}\Big[\sum_{j\neq i} b_j\,\alpha_j(\ell^*, \sigma)\,q^*_j(b, \sigma)\Big]. \tag{1}$$

Let $s_i(\sigma) := \max_{j\neq i} b_j\alpha_j(\ell^*, \sigma)$. For every realized σ, the welfare-maximizing $q^*(b, \sigma)$ allocates only to maximizers of $b_j\alpha_j(\ell^*, \sigma)$. Therefore, for every signal σ

$$\sum_{j\neq i} b_j\alpha_j(\ell^*, \sigma)\,q^*_j(b, \sigma) = s_i(\sigma)\sum_{j\neq i} q^*_j(b, \sigma).$$

If $s_i(\sigma) > 0$, since $\sum_j q_j^*(b,\sigma) = 1$, we have $\sum_{j\neq i} q_j^*(b,\sigma) = 1 - q_i^*(b,\sigma)$. If $s_i(\sigma) = 0$, both sides are 0 regardless of q^*. Therefore, for all σ,

$$\sum_{j\neq i} b_j \alpha_j(\ell^*,\sigma)\, q_j^*(b,\sigma) = s_i(\sigma)\left(1 - q_i^*(b,\sigma)\right). \tag{2}$$

Taking an expectation in (2) and substituting into (1) gives

$$\begin{aligned} p_i^{VCG}(b) &= \max_\ell \mathrm{SW}_{-i}(\ell) - E_{\sigma\sim S_{\ell^*}}[s_i(\sigma)] + E_{\sigma\sim S_{\ell^*}}[s_i(\sigma)\, q_i^*(b,\sigma)] \\ &= \underbrace{\left(\max_\ell \mathrm{SW}_{-i}(\ell) - \mathrm{SW}_{-i}(\ell^*)\right)}_{\text{Stage–1 externality}} + \underbrace{E_{\sigma\sim S_{\ell^*}}\left[s_i(\sigma)\, q_i^*(b,\sigma)\right]}_{\text{expected Stage–2 second price}}. \end{aligned}$$

□

Proof of Theorem 4.1. For any realized (ℓ,σ), $\mathcal{M}_A^{VCG}$ is a single-item VCG auction on effective values $b_i^A(\ell,\sigma)\cdot\alpha_i(\ell,\sigma)$. Hence, $b_i^A(\ell,\sigma) = v_i$ is a (weakly) dominant strategy for every advertiser i, independent of stage 1 behavior. Therefore, any strategy s_i containing a bidding function $b_i^A(\ell,\sigma) \neq v_i$ is weakly dominated by a strategy s_i' that is identical to s_i but sets $b_i^A(\ell,\sigma) = v_i$. We thus restrict our attention to strategies where $b_i^A(\ell,\sigma) = v_i$.

Let $u_i^A(v;\ell,\sigma)$ denote i's quasilinear utility in stage 2 outcome (fixing $b_i^A(\ell,\sigma) = v_i$ for all i,ℓ,σ), and let $R_i(\ell;v) := E_{\sigma\sim S_\ell}\left[u_i^A\big(v;\ell,\sigma\big)\right]$ be each advertiser's ex-ante utility for implementing question ℓ, fixing the behavior in stage 2. Advertiser i's total expected utility under stage 1 bids b^Q is

$$U_i(b^Q) \;=\; \sum_{\ell=1}^{k} \lambda_\ell(b^Q)\, R_i(\ell;v) \;-\; p_i^Q(b^Q).$$

The prescribed strategy in stage 1 is $b_{i,\ell}^Q = R_i(\ell;v)$, for every advertiser i and question ℓ. Fix the bids of all other advertisers j to $b_{j,\ell}^Q = R_j(\ell;v)$, and consider an alternative stage 1 bidding strategy $\hat{b}_{i,\ell}^Q$. Slightly overloading notation, it remains to argue that $U_i(R_i;b_{-i}) \geq U_i(\hat{b}_i^Q;b_{-i})$.

First, if the chosen question under $(\hat{b}_i^Q,b_{-i})$ is the same as the chosen question under $(R_i;R_{-i})$, then the stage 1 payment $p_i^Q(\hat{b}_i^Q,b_{-i})$ is equal to $p_i^Q(R_i,b_{-i})$ (since the externality is the same), so $U_i(\hat{b}_i^Q;b_{-i}) = U_i(R_i;b_{-i})$. Second, consider the case that the chosen question $\hat{\ell}$ under $(\hat{b}_i^Q,b_{-i})$ is *not* the same as the chosen question ℓ^* under $(R_i;b_{-i})$. Then we have

$$\begin{aligned} U_i(\hat{b}_i^Q;b_{-i}) &= R_i(\hat{\ell};v) - p_i^Q(\hat{b}_i^Q,b_{-i}) \\ &= \sum_{j=1}^{n} R_j(\hat{\ell};v) - \max_\ell \sum_{j\neq i} R_j(\ell;v). \end{aligned}$$

The second term does not depend on i. And, the first term, $\sum_{j=1}^n R_j(\hat{\ell};v)$ is maximized for $\hat{\ell} = \ell^*$. Therefore,

$$\begin{aligned} U_i(\hat{b}_i^Q;b_{-i}) &= \sum_{j=1}^{n} R_j(\hat{\ell};v) - \max_\ell \sum_{j\neq i} R_j(\ell;v) \\ &\leq \sum_{j=1}^{n} R_j(\ell^*;v) - \max_\ell \sum_{j\neq i} R_j(\ell;v) \\ &= U_i(R_i;b_{-i}). \end{aligned}$$

This concludes the proof that the prescribed strategy is an undominated pure Nash equilibrium. □

Proof of Theorem 4.3. For first-price and all-pay, we use the instance from the proof of Theorem 4.2. We begin by noting that the second-stage mechanism is still a VCG auction. We therefore assume that, in the equilibrium we analyze, advertisers will play their dominant strategy and bid their true base values ($b_i^A(\ell,\sigma) = v_i$) in stage 2. This allows us to use the derived expected utilities, $R_i(\ell;v)$, as the advertisers' values in the stage 1 auction:

- $R_i(1; v) = \delta/m$ for all advertisers i.
- $R_1(2; v) = \frac{1-\delta}{m}$, and $R_i(2; v) = 0$ for all $i \neq 1$.

Now, consider a first-price auction (the advertiser with the highest bid on the winning question pays this bid) or all-pay auction (all advertisers pay their bid on the winning question) for stage 1, where the question with the highest sum of bids wins. We will assume tie-breaking in favor of question 2: that is, if the sum of bids is the same, then question 2 (the uninformative question) is chosen. The instance and proof can be easily adjusted to work under tie-breaking that favors the first question.

We claim the following strategy profile for stage 1 is a Nash Equilibrium for conservative bidders (as defined in Definition 4.1). Advertiser 1 bids $b^Q_{1,2} = \frac{(m-1)\delta}{m}$ for the uninformative question $\ell = 2$ and $b^Q_{1,1} = 0$. All other advertisers $i \neq 1$ bid $b^Q_{i,2} = 0$ for the uninformative question, and $b^Q_{i,1} = \frac{\delta}{m}$ for the informative question. Note that all bids $b^Q_{j,\ell}$ are at most $v_j E_{\theta\sim\mathcal{D}}[\alpha_j(\theta)]$ (which is $\frac{3-2\delta}{m}$ for advertiser 1, $\frac{1}{m}$ for advertiser 2, and $\frac{2-\delta}{m}$ for all other advertisers, by picking δ sufficiently small).

In this profile, the uninformative question $\ell = 2$ is chosen (the sum of bids on both questions is $\frac{(m-1)\delta}{m}$; by our tie-breaking rule, question 2 wins), yielding a welfare of $\frac{3-2\delta}{m}$, while the optimal welfare is 1. It remains to argue that this is an equilibrium.

First, advertiser 1 does not want to increase its bid for the winning question, since this action would increase its payment without changing the outcome. Increasing its bid for the losing question will change the outcome to question 1 being picked, which gives utility at most $\frac{\delta}{m}$, which is strictly less than her current utility of $R_1(2; v) - \frac{(m-1)\delta}{m} = \frac{1-\delta}{m} - \frac{(m-1)\delta}{m}$, for all $\delta < \frac{1}{m+1}$. Advertiser 1 cannot decrease its bid for the losing question (it's already zero). Finally, advertiser 1 does not want to decrease its bid for the winning question: any decrease of $\epsilon > 0$ would result in a different question being asked. Currently, advertiser 1 gets utility $R_1(2; v) - \frac{(m-1)\delta}{m} = \frac{1-\delta}{m} - \frac{(m-1)\delta}{m}$ which is at least $R_1(1; v) = \frac{\delta}{m}$ for all $\delta < \frac{1}{m+1}$.

Second, the only thing that an advertiser $i \neq 1$ can do to affect the outcome is to bid strictly more than $\frac{\delta}{m}$ on question 1. However, $R_i(1; v) = \frac{\delta}{m}$, so, since under both payment rules (first-price and all-pay) they will end up paying their bid, this deviation is not profitable. Finally, if advertiser i's deviation, for $i \neq 1$, does not affect the outcome (question asked), then its payment cannot be further reduced (since its bidding zero for the winning question already). □

Proof of Lemma 4.1. Since both VCG and the Bailey-Cavallo mechanism have dominant strategies, under our rationality assumption it suffices to consider the corresponding equilibrium.

First, in the Bailey-Cavallo redistribution mechanism, it is a dominant strategy for each advertiser i to bid their true base value, i.e., $b^A_i = v_i$ for all i. Therefore, for every equilibrium in undominated strategies, the second stage allocation is efficient, generating optimal welfare $W^*(\ell, \sigma) = \max_i v_i \alpha_i(\ell, \sigma)$. Furthermore, in the Bailey-Cavallo mechanism the revenue is at most $\frac{2}{n} W^*(\ell, \sigma)$, and therefore the sum of advertisers' utilities is at least $(1 - \frac{2}{n}) W^*(\ell, \sigma)$. Taking expectations, we have that $\sum_{i=1}^n R_i(\ell) \geq \left(1 - \frac{2}{n}\right) W(\ell)$, where $R_i(\ell) = E_{\sigma\sim S_\ell}[u_i(\ell, \sigma)]$ is the expected utility of advertiser i for question ℓ, and $W(\ell) = E_{\sigma\sim S_\ell}[W^*(\ell, \sigma)]$ is the expected optimal welfare.

Now consider the first stage. Since the mechanism is a VCG auction, it is a dominant strategy for each advertiser to bid their true expected value on each question; here, the value for question ℓ is the expected stage 2 utility, $R_i(\ell)$. Thus, in an undominated equilibrium, advertisers submit bids $b^Q_{i,\ell} = R_i(\ell)$, and the mechanism selects the question ℓ^* that maximizes the sum of reported bids, i.e., $\ell^* \in \operatorname{argmax}_\ell \sum_{i=1}^n R_i(\ell)$.

Let ℓ_{opt} be the question that maximizes the expected social welfare (i.e., $W(\ell_{opt}) \geq W(\ell)$ for all ℓ). We have:

$$W(\ell^*) \geq \sum_{i=1}^n R_i(\ell^*) \geq \sum_{i=1}^n R_i(\ell_{opt}) \geq \left(1 - \frac{2}{n}\right) W(\ell_{opt}).$$

Rearranging this inequality yields the Price of Anarchy bound: $\frac{W(\ell_{opt})}{W(\ell^*)} \leq \frac{1}{1-\frac{2}{n}} = 1 + \frac{2}{n-2}$. □

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: All proofs are included, either in the main body or appendix.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: We discuss limitations of the model in the intro.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: The formal model is presented in section 2. All proofs are included in the main body and appendix.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [N/A]

   Justification: The paper does not include experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [N/A]

Justification: The paper does not include experiments.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [N/A]

   Justification: The paper does not include experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

   Justification: The paper does not include experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)

- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A]

   Justification: There are no experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: We have reviewed, and conform to the code of ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: This is a theoretical paper, on foundational mechanism design questions. There is limited societal impact of the work performed. We briefly comment on this in page 2, footnote 1.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.

- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: The paper poses no such risks.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [N/A]

    Justification: The paper does not use existing assets.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.

- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: The paper does not release new assets.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A]

Justification: The core method development in this research does not involve LLMs as any important, original, or non-standard components. We have used LLMs only for editing/help with word-choice.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.