DATA-DRIVEN MECHANISM DESIGN: JOINTLY ELICITING PREFERENCES AND INFORMATION

By
Dirk Bergemann, Marek Bojko, Paul Dütting,
Renato Paes Leme, Haifeng Xu and Song Zuo

December 2024

COWLES FOUNDATION DISCUSSION PAPER NO. 2418

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Data-Driven Mechanism Design: Jointly Eliciting Preferences and Information

Dirk Bergemann* Marek Bojko† Paul Dütting‡

Renato Paes Leme§ Haifeng Xu¶ Song Zuo‖

December 20, 2024

Abstract

We study mechanism design when agents hold private information about both their preferences and a common payoff-relevant state. We show that standard message-driven mechanisms cannot implement socially efficient allocations when agents have multidimensional types, even under favorable conditions.

To overcome this limitation, we propose data-driven mechanisms that leverage additional post-allocation information, modeled as an estimator of the pay-off relevant state. Our data-driven mechanisms extend the classic Vickrey-Clarke-Groves class. We show that they achieve exact implementation in posterior equilibrium when the state is either fully revealed or the utility is linear in an unbiased estimator. We also show that they achieve approximate implementation with a consistent estimator, converging to exact implementation as the estimator converges, and present bounds on the convergence rate. We demonstrate applications to digital advertising auctions and large language model (llm) - based mechanisms, where user engagement naturally reveals relevant information.

Keywords: Mechanism Design, Data-Driven Mechanism Design, Large Language Models, Click-Through Rate, Posterior Equilibrium

JEL Codes: D47, D82, D83

*Department of Economics, Yale University, dirk.bergemann@yale.edu

†Department of Economics, Yale University, marek.bojko@yale.edu

‡Google Research, duetting@google.com

§Google Research, renatoppl@google.com

¶Department of Computer Science, The University of Chicago and Google Research, haifengustc@google.com

‖Google Research, szuo@google.com

*Dirk Bergemann gratefully acknowledges financial support from NSF SES 1948336 and SES 2049754 and ONR MURI. We thank participants of the Yale Micro-Theory Breakfast seminar for valuable comments.

1 Introduction

1.1 Motivation

The rise of data-rich digital environments has created new opportunities and challenges for mechanism design. In settings ranging from online advertising to large language models (LLM), participants often possess private information not just about their preferences but also about underlying states that affect all agents' payoffs. For instance, in sponsored search auctions, advertisers have private information about both their value per click and user behavior patterns. In emerging LLM applications, content providers have private knowledge of both their desired outcomes and user preferences.[1] Efficiently aggregating and incentivizing the revelation of this multidimensional private information is crucial for the optimal allocation. While the tension between disentangling preferences from information in mechanism design and related environments is well-documented,[2] these "modern" applications highlight the timeliness and relevance of revisiting these questions. At the same time, these applications provide an abundance of additional data revealed through user interactions, platform feedback, or other observable outcomes.[3] As we will show, this additional data can be used to reconcile these two forces.

In a static setting with interdependent values and quasi-linear preferences, we develop a framework for mechanism design that moves beyond traditional message-driven approaches by incorporating such naturally available data. Rather than relying solely on agents' reported messages, our mechanisms condition transfers on additional information about the state. This data-driven approach enables the implementation of efficient allocations, maximizing the sum of agents' expected payoffs based on their combined information, even when agents have multidimensional private types—a setting where standard mechanisms provably fail. We specifically focus on implementation in posterior equilibrium, ensuring that no agent is incentivized to deviate from truthful reporting, provided that all other agents report truthfully, even after the uncertainty regarding other agents' types has resolved. Implementation in posterior equilibrium thus ensures robustness to belief misspecification and eliminates incentives for agents to inefficiently acquire information about other agents' types. As a result, it promotes stability and strategic simplicity in complex environments.

1.2 Results

We begin our analysis with the standard message-driven mechanism design framework, where allocations and transfers depend solely on the messages sent by agents to the mechanism. In in-

[1] A prominent example is the novel format of digital advertising auctions, where advertisers bid to feature in "sponsored" content generated by an LLM. At the time of this writing, Perplexity AI, a widely used LLM-based chatbot with an integrated search engine, began incorporating such digital advertising auctions on its platform; see this Financial Times article: https://on.ft.com/47BWbIX.

[2] See, for example, Lu (2019) and references therein for a related issue of identification of beliefs and state-dependent utilities.

[3] For instance, in sponsored search auctions, this additional information includes click-through rates on ads. In settings involving LLMs, it encompasses metrics such as clicks on suggested links, follow-up queries, and engagement duration.

terdependent value settings, Maskin (1992), Dasgupta and Maskin (2000), Jehiel and Moldovanu (2001), and Bergemann and Välimäki (2002) established an implementation impossibility when agents have multidimensional types or, in the case of single-dimensional types, when an appropriate single-crossing condition is unmet. Given that agents in our model possess private information regarding both the state and their preferences, private types are inherently multidimensional. In Theorem 1, we show that implementation in our setting remains impossible even under conditions where previous possibility results emerged: there is an instance with single-dimensional signals and supermodular expected payoffs for which no message-driven mechanism implements the efficient allocation in posterior equilibrium.

To address the impossibility result, we adopt an approach similar to Mezzetti (2004) by relaxing the restriction that transfers rely solely on the agents' messages. Instead, we allow transfers to be conditioned on additional information about the state, while maintaining that allocations are determined only by messages. In a setting similar to ours, Mezzetti (2004) proposed a two-stage mechanism: in the first stage, the mechanism elicits agents' types to determine the allocation, and in the second stage, it elicits realized payoffs to compute transfers. However, this approach assumes that agents can fully observe and report their final payoffs, and that the second stage is feasible—conditions that may be restrictive or unworkable due to institutional constraints.[4] We show implementation in posterior equilibrium can be obtained under strictly weaker requirements.

We formalize the additional information as a state estimator available to the designer after allocation but before finalizing transfers, with the estimator's data-generating process being common knowledge. In our motivating examples, this approach captures scenarios where the designer gathers data on user interactions both within and beyond the auction environment, thereby gaining insights into user preferences and demand (i.e., the state). As highlighted, such data is naturally available and already collected by platforms, providing a straightforward interpretation for our framework. Notably, unlike Mezzetti (2004), our method eliminates the need for a second reporting stage or access to agents' true preferences, thereby reducing the amount of information necessary for implementation.

We introduce a modified version of Vickrey–Clarke–Groves (VCG) mechanisms (Vickrey, 1961; Clarke, 1971; Groves, 1973), which we call data-driven VCG mechanisms (Definition 6). In data-driven VCG transfers, the designer plugs in the estimator into agents' payoffs. We then examine how various properties of the estimator affect the feasibility of implementation. We first show that with full resolution of uncertainty, data-driven VCG mechanisms achieve implementation in posterior equilibrium (Proposition 1). Specifically, if the designer ultimately observes the true state, she can utilize ex-post utilities based on the reported types to calculate VCG payments, aligning agents' incentives. This alignment occurs because agents assess these

[4]For example, Jehiel and Moldovanu (2005) highlight the fragility of mechanisms proposed by Mezzetti (2004). Specifically, agents may have only noisy and subjective perceptions of the state, which affect both their payoffs and the reliability of the mechanism. This fragility is further compounded by issues such as moral hazard, verifiability, and other practical concerns. Furthermore, in the second stage of Mezzetti (2004)'s two-stage generalized VCG mechanism, agents are indifferent to their utility reports. This contrasts with the typical strict incentives for truthful reporting in standard VCG mechanisms, particularly in settings with private values.

transfers according to their true posterior beliefs, making truthful reporting optimal by the efficiency of the allocation, as per a standard argument.

Next, we examine the scenario where the designer only has access to a noisy estimate of the state and assess the implications of two key properties of estimators common in statistics and econometrics. First, when using an unbiased estimator and assuming agents' utility functions are linear in the state, data-driven VCG mechanisms yield implementation in posterior equilibrium (Proposition 2). However, as anticipated by Jensen's inequality, this result does not hold universally beyond linear utilities. Nevertheless, our application to click-through auctions demonstrates that an unbiased estimator can be highly beneficial in settings with risk-neutral bidders and appropriately defined states.

Second, we consider the property of consistency, where a sequence of estimators, indexed by the dataset size, converges in probability to the true state as the dataset size approaches infinity. Subject to regularity conditions on agents' utility functions, Theorem 2 establishes that any corresponding sequence of data-driven VCG mechanisms achieves implementation in ϵ-posterior equilibrium, where no agent has more than an ϵ additive regret of having reported truthfully after the resolution of uncertainty about others' types when other agents also report truthfully, with ϵ approaching zero as the sample size increases. That is, while reporting truthfully is an ϵ-posterior equilibrium in any finite sample for large enough ϵ for any estimator, consistent estimators ensure the ϵ can be made arbitrarily small in the limit. Proposition 3 further extends this result by linking the rate of convergence of the estimator to the rate at which ϵ approaches zero. Assuming both regularity in agents' utility functions and uniform integrability of the sequence of relevant random variables, we demonstrate that the sequence of ϵ's can be reduced at essentially the same rate as the estimator converges in probability to the true state.

1.3 Applications

We first apply our framework and results to click-through auctions, focusing on a position auction for a single advertising slot for simplicity. Following the standard specification of agents' utility functions (Edelman et al., 2007; Varian, 2007), we treat agents' values per click as their preference types and the click-through rates as the state. Within this framework, mechanisms based solely on per-impression payments—conditioned only on agents' actions—are subject to the same impossibility result identified earlier. In contrast, the commonly used per-click payments fall within the class of data-driven mechanisms. However, their effectiveness in eliciting truthful reporting of both agents' values and signals hinges on whether the click-through rates are common or agent-specific.

When click-through rates are common, per-click pivot payments achieve implementation in posterior equilibrium. These payments align with the data-driven VCG mechanisms class, as the observed frequency of clicks serves as an unbiased (and consistent) estimator of the true click-through rate if clicks are sampled independently and identically (i.i.d.). However, with agent-specific click-through rates, data-driven VCG transfers cannot be feasibly structured as per-click payments, and feasible per-click payment rules become vulnerable to manipulation.

Our second application focuses on mechanism design for content generated by LLMs. Each input prompt defines a distinct mechanism design environment. We model an LLM as a conditional prediction system that selects a generation distribution over feasible output text units to maximize the expected value of a provided reward function, leveraging its training data, knowledge bases, and model architecture—represented as signals and preference types within our framework. We introduce a reference LLM that represents the platform's organic generation distribution, aiming to align responses with user preferences. The designer's objective is to determine a central LLM generation distribution that maximizes the total reward of agents while accounting for a regularization term that measures deviations from the reference distribution. Incorporating this regularization term into each agent's transfer within data-driven VCG mechanisms leads to the definition of regularized data-driven VCG mechanisms (Definition 9). These mechanisms retain the key properties established in our main analysis (Corollary 3). To illustrate, we provide an example where the regularization term is specified as the Kullback-Leibler divergence between the selected distribution and the reference distribution (Example 4).

1.4 Related Literature

This paper contributes to several strands of the mechanism design literature.

Efficient mechanism design with interdependent values. The literature on mechanism design with interdependent values, initiated by Milgrom and Weber (1982) and expanded in subsequent work, is extensive. More recent research (Dasgupta and Maskin, 2000; Jehiel and Moldovanu, 2001; Bergemann and Välimäki, 2002) has shown that incentive compatibility and efficiency cannot be simultaneously achieved when signals are independent and multidimensional, or single-dimensional if the single-crossing condition is violated. These impossibility results extend to implementation in posterior equilibrium, irrespective of the stochastic structure of the signals.

Our impossibility result builds on this body of work, relying on the assumption that agents have a strictly positive informational size, as defined by McLean and Postlewaite (2002). In contrast, McLean and Postlewaite (2015) demonstrate that when agents have zero informational size, implementation in posterior equilibrium is possible. Furthermore, they establish a continuity result: as agents' informational size approaches zero, approximate posterior equilibrium implementation becomes feasible, converging to exact implementation in the limit. We also provide a continuity result, albeit through a different channel, while allowing for arbitrary informational sizes of the agents.

Several papers propose two-stage mechanisms to align incentives. In McLean and Postlewaite (2017)'s mechanism, agents first report their signals. The designer collects the information and announces the implied posterior on the state. This reduces the problem to a private value setting in the second stage, where agents report their expected payoffs for each social outcome. Allocations and transfers are then determined using a VCG mechanism based on these reports. Mezzetti (2004) considers a framework where the allocation is fixed upfront, but transfers depend on agents' subsequent reports of their final payoffs. The central result establishes that efficient allocation can always be implemented in Perfect Bayesian Equilibrium via a two-stage VCG mechanism. Instead, we show how additional information often readily available to the designer can be utilized to align incentives without requiring further communication. Our data-driven transfer scheme also addresses practical challenges, such as when agents cannot observe final payoffs or when a two-stage mechanism is infeasible. Furthermore, compared to Mezzetti (2004), our results imply that implementation is achievable with strictly less information transmitted to the mechanism. Specifically, agents' true ex-post payoffs need not be reported for implementation in posterior equilibrium.

Wu et al. (2024) study an auction design setting where bidders also have both preference and information. However, for information component, they only allow bidders to reveal partial information but forbid misreports, whereas our model allows information misreporting.

Mechanism design with contingent payments and data-driven components. Our data-driven transfer scheme also connects with the literature on mechanisms involving contingent payments and public ex-post information (e.g. the work of Hansen (1985) and Riordan and Sappington (1988), and the follow-on literature). Additionally, a growing body of research examines how data on agent types can facilitate efficient outcomes in environments with adverse selection (Braverman and Chassang, 2022; She et al., 2022; Liang and Madsen, 2024). Our focus diverges by exploring how a designer can leverage data on a common, payoff-relevant stochastic factor to design transfers that encourage truthful reporting of agents' types.

Sponsored-search auctions. The standard model for sponsored-search auctions due to Varian (2007) and Edelman et al. (2007) assumes that bidders have a value per click and ad slots have non-increasing click-through rates. Both studies independently propose (essentially the same) refinement of (pure) Nash equilibrium, and argue that under this equilibrium concept the Generalized-Second Price (GSP) mechanism yields an efficient allocation. In fact, as shown by Caragiannis et al. (2015), even worst-case equilibria of the GSP mechanism are close to optimum for a broad range of equilibrium concepts. A particular feature of the GSP mechanism is that it can be implemented without knowledge of the click-through rates. This is also true for the VCG mechanism, but requires a more careful implementation in which bidders get charged and credited (Varian, 2009; Varian and Harris, 2014).

Milgrom (2010) argues that the common practice of asking advertisers for a single bid (their value per click) can be beneficial (e.g., to rule out low-revenue equilibria). However, he also points out the associated risk of model mis-specification. Specifically, he argues that there may be situations where clicks are perfectly observable, but an advertiser's value may be misaligned with observable clicks. As a concrete example, he mentions a situation where there are two types of searchers—potential shoppers who are actually looking to buy a product and others that are merely curious. The two groups may have different click-through rates, but only clicks from potential shoppers are valuable. Motivated by this, Dütting et al. (2019) and Dütting et al. (2024) examine standard position auctions with mis-specified bidding languages; and establish a ranking between standard position auction formats in regard to the format's ability to support an efficient equilibrium, when values follow different click-through rates than those used in the auction. Our work approaches this differently, by modeling the click-through rates as private information that the bidders can report to the auctioneer.

Another related line of work is Bergemann et al. (2022) and Chen et al. (2023), who consider a situation where click-through rates are stochastic; and the auctioneer strategically discloses information about the distribution of click-through rates so as to maximize revenue. In this model the auctioneer has additional information, which is in contrast to our model where it is the bidders that hold additional information.

Mechanism design for LLMs. The nascent strand of literature on mechanism design for LLM-based auctions began with Dütting et al. (2024), who proposed a sequential auction mechanism for creative ad generation where advertisers bid on a token-by-token basis. In this framework, each advertiser uses its own LLM to generate ad content, which is assumed to be truthfully reported to a central LLM. The token auction takes single-dimensional bids as inputs and produces a joint output by aggregating individual LLM distributions. They also develop an analogue of the second-price auction for this setting, demonstrating desirable incentive properties. Soumalias et al. (2024) propose an auction that allows agents to influence the output of the central LLM according to their reported reward functions, with an aggregate cost based on the distance between the generated LLM and a reference LLM. They design a sampling allocation scheme and payment rule to ensure incentive compatibility. Meanwhile, Dubey et al. (2024) examine a setting where advertisers bid for prominence in an LLM-generated summary, constructing an auction to allocate ad prominence within the output. Similarly, Hajiaghayi et al. (2024) propose an auction mechanism that probabilistically selects the ad displayed in each segment following the Retrieval-Augmented Generation framework.

In contrast to these approaches, we examine a setting where agents not only have preferences over the central LLM's output but also possess valuable information that can improve its accuracy. Additionally, we explore a richer environment in terms of potential manipulations and mechanism design. For example, unlike Soumalias et al. (2024), our framework requires agents to pay for introducing bias into the final allocation, with the payment increasing according to the marginal impact of the introduced bias.

1.5 Organization of the Paper

The paper is organized as follows. Section 2 introduces the formal model. In Section 3, we analyze the implementation problem in the setting of standard message-driven mechanisms, establishing an impossibility result for implementation in posterior equilibrium. In Section 4, we define data-driven mechanisms, introduce the data-driven VCG mechanism, and present our implementation results within this framework. Sections 5.1 and 5.2 apply these results to click-through auctions and LLM-based environments, respectively. Finally, Section 6 discusses the implications of our model, summarizes the key findings, proposes directions for future research, and concludes the paper.

2 Model

We consider the following mechanism design setting. There is a set $N \equiv \{1, \ldots, n\}$ of agents. Let X be the space of feasible allocations. Uncertainty is represented by a set of possible states of the world Ω, with a typical element ω. X and Ω are assumed to be compact subsets of Euclidean spaces. We denote by $\|\cdot\|$ the Euclidean norm unless otherwise specified.

Each agent i has private information about her *preference type* $\theta_i \in \Theta_i$ and a *private signal* $s_i \in \mathcal{S}_i$. Tuple (θ_i, s_i) forms agent i's *type*. Define $\Theta = \prod_{i=1}^n \Theta_i$ and $\mathcal{S} = \prod_{i=1}^n \mathcal{S}_i$, with generic elements $\theta = (\theta_1, \ldots, \theta_n)$ and $s = (s_1, \ldots, s_n)$. Θ_i and $\mathcal{S}_i$ are assumed to be compact subsets of Euclidean spaces for each agent i.

There is a full-support prior $P \in \Delta(\Omega \times \Theta \times \mathcal{S})$ on the state and agents' preference types and signals which is common knowledge. We assume s_i is sufficient for ω with respect to θ_i: $P(\omega|s_i, \theta_i) = P(\omega|s_i)$ for each agent i, $\theta_i \in \Theta$, and $s_i \in \mathcal{S}_i$. Observing a profile of signal realizations s leads each agent to update her prior belief on ω according to Bayes' rule, with a posterior distribution $P(\omega|s)$.

Agent i's utility is quasi-linear in the payoff and transfers: $U_i : X \times \Omega \times \Theta_i \times \mathbb{R} \to \mathbb{R}$ given by:

$$U_i(x, \omega, \theta_i, t_i) = u_i(x, \omega, \theta_i) + t_i,$$

where we assume u_i is common knowledge and continuous. Agent i's expected payoff from decision x, conditional on θ_i and s, is given by:[5]

$$v_i(x, \theta_i, s) \equiv \int_\Omega u_i(x, \omega, \theta_i) dP(\omega|s).$$

An *instance* is a tuple $\Gamma = (N, X, \Omega, \Theta, \mathcal{S}, P, (U_i)_{i \in N})$.

This paper focuses on implementing the efficient allocation rule based on the combined information held by agents while allowing for residual uncertainty about the state of the world.

Definition 1. (Efficient allocation rule)
The deterministic allocation rule $x : \Theta \times \mathcal{S} \to X$ is efficient if, for all $\theta \in \Theta$ and $s \in \mathcal{S}$, it satisfies

$$x(\theta, s) \in \arg\max_{x \in X} \sum_{i \in N} v_i(x, \theta_i, s). \tag{1}$$

We will denote a generic efficient allocation rule by x^*. Note that it is always well-defined by our continuity and compactness assumptions.

3 Message-Driven Mechanisms

In a *message-driven mechanism*, agents are asked to select from a menu of messages and the allocation and transfers are determined solely based on the agents' selection. This is, of course, a standard notion in the mechanism design literature. It is without loss of generality

[5]Similarly to McLean and Postlewaite (2015), our formulation of agents' utilities highlights the source and type of interdependence: agents care about others' signals only in so far as they provide valuable information about the state.

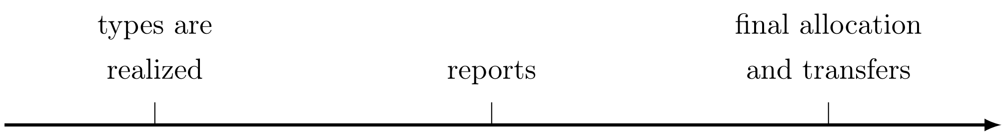

Figure 1: Timing of a message-driven direct revelation mechanism (Definition 2).

to focus on direct revelation mechanisms by the revelation principle. The standard timeline of these mechanisms is illustrated in Figure 1.

Definition 2. (Message-driven direct revelation mechanism)
A *message-driven direct revelation mechanism* is a pair (x, t) where $x : \Theta \times \mathcal{S} \to X$ is the outcome function and $t : \Theta \times \mathcal{S} \to \mathbb{R}^N$ are the transfers.

We focus on efficient implementation in posterior equilibrium, an equilibrium notion first defined by Green and Laffont (1987).

Definition 3. (Posterior equilibrium with message-driven mechanisms)
A message-driven direct revelation mechanism (x, t) *permits implementation in posterior equilibrium* if for each agent $i \in N$ and true types $\theta \in \Theta$ and $s \in \mathcal{S}$,

$$v_i(x(\theta, s), \theta_i, s) + t_i(\theta, s) \geq v_i(x(\theta'_i, \theta_{-i}, s'_i, s_{-i}), \theta_i, s) + t_i(\theta'_i, \theta_{-i}, s'_i, s_{-i}), \forall s'_i \in \mathcal{S}_i, \theta'_i \in \Theta_i.$$

In essence, once uncertainty about other agents' types is resolved, no agent has any regret about reporting her type truthfully. That is, truthful reporting remains optimal for each agent given that others also report truthfully—ex-post with respect to agent types (though not necessarily the state).[6] This equilibrium concept is appealing because it does not require agents to specify beliefs over others' types and ensures they will not wish to alter their actions after the allocation is finalized.

Maskin (1992), Dasgupta and Maskin (2000), Jehiel and Moldovanu (2001), and Bergemann and Välimäki (2002) demonstrated the impossibility of implementation with interdependent values when types are either multidimensional or single-dimensional without a suitable single-crossing condition.[7] In our setting, even with single-dimensional signals and a single-crossing condition, types remain multidimensional. We obtain the following result.

Theorem 1. *(Impossibility)*
There is an instance where the signal of each agent is single-dimensional and v_i is supermodular

[6]Note that the posterior equilibrium concept defined by Green and Laffont (1987) for an arbitrary mechanism requires agents' strategies to be optimal against each other's strategies, based on the specific information revealed by the mechanism. In a direct revelation mechanism, this corresponds to revealing the transmitted messages, i.e., the reported types, leading to the notion defined here. The idea that agents have no regret about truthful reporting after uncertainty about others' types is resolved is often referred to as *ex-post equilibrium* (Bergemann and Välimäki, 2002). While these two concepts coincide in our context, they generally differ (Jehiel et al., 2007). Given the importance of resolving uncertainty about the state in our setting, we adopt the terminology of *posterior* equilibrium for discussing implementation.

[7]Each of these works assumes independent signals, but the results extend to cases of arbitrarily correlated signals when considering implementation in posterior equilibrium.

in (x, s_j) for each i and j but for which no message-driven mechanism implements the efficient allocation rule in posterior equilibrium.

To prove the impossibility result, we provide the following example, which will also serve as our running example later in the paper.

Example 1. (Quadratic loss: impossibility)
There are two agents with single-dimensional preference types, where we assume $\Theta_1 = \Theta_2 \subset \mathbb{R}$ are closed intervals sufficiently wide to make the objects below well-defined, and signals, $\mathcal{S}_1 = \mathcal{S}_2 \subseteq \mathbb{R}$. Agent i's payoff is given by

$$u_i(x, \omega, \theta_i) = -(x - \theta_i - \omega)^2.$$

The efficient allocation is then given by

$$x^*(\theta, s) = \frac{\theta_1 + \theta_2}{2} + \mathbb{E}_{\omega|s_1,s_2}[\omega].$$

That is, θ_i can be interpreted as the *bias* agent i seeks to introduce into the allocation.

We assume the following condition on the informativeness of signals.[8] There are signals $s_1, s_1' \in \mathcal{S}_1, s_1' \neq s_1$ and $s_2 \in \mathcal{S}_2$ such that

$$\mathbb{E}_{\omega|s_1,s_2}[\omega] \neq \mathbb{E}_{\omega|s_1',s_2}[\omega]. \tag{$*$}$$

We also assume the mapping $x \mapsto \mathbb{E}_{\omega|x,s_{-i}}[\omega]$ is differentiable for each i and s_{-i} and increasing. These conditions are satisfied, for example, if $\omega \sim N(\mu, \sigma^2)$ with unknown μ and for each i, $s_i = \omega + \epsilon_i$ with independent $\epsilon_i \sim N(0, \sigma_i^2)$ and known variances σ_i^2.[9] Under these assumptions, $\frac{\partial v_i(x,\theta_i,s)}{\partial x \partial s_j} \geq 0$ for any $i, j \in N, x \in X, \theta_i \in \Theta_i$, and $s \in \mathcal{S}$.

We now prove that no message-driven mechanism can implement the efficient allocation rule in posterior equilibrium. By the revelation principle, it suffices to demonstrate this for direct revelation mechanisms. We proceed by contradiction. Suppose there exist s_1, s_1', s_2 satisfying $(*)$ and a transfer rule t that depends only on agents' reports such that (x^*, t) implements the efficient allocation in posterior equilibrium. Then t must prevent each agent from having a profitable deviation by misreporting either her preference type or her signals.

Using standard results from the literature, we characterize the class of transfer schemes that deter profitable deviations along each dimension of agents' types. Specifically, if only preference types are private and signals are truthfully reported, the transfer scheme must belong to the class of standard VCG payments. Conversely, if signals are private and reported truthfully, the transfer scheme must follow the class of generalized VCG payments described by Bergemann and Välimäki (2002). The fact that these two classes of payment rules are generally distinct helps us reach the desired contradiction. Details of the argument are provided below.

[8] Note that the condition is satisfied if the signals of all agents jointly determine the state; for example, in the wallet model: $\omega = \sum_{i \in N} s_i$ (Klemperer, 1998). The result thus directly applies to this benchmark case.

[9] To ensure these distributional assumptions match the current framework where all spaces considered are compact, we can suitable truncate these distributions. As standard algebra reveals, the formula for the posterior mean and variance is essentially unchanged.

Suppose the true profile of signals is $s = (s_1, s_2)$ and it is fixed to be reported truthfully by both agents. Consider the class of VCG transfers for agent 1:

$$\begin{aligned} t_1(\theta, s; h_1) &= h_1(\theta_2; s) + \mathbb{E}_{\omega|s}\Big[u_2\big(x^*(\theta, s), \omega, \theta_2\big)\Big] \\ &= h_1(\theta_2; s) - \frac{1}{4}(\theta_1 - \theta_2)^2 - \mathrm{Var}_{\omega|s}[\omega], \end{aligned}$$

for an arbitrary function $h_1(\theta_2; s)$ of θ_2. We can define the class of VCG transfers for agent 2 analogously. As is well-known, the VCG mechanism implements the efficient outcome in posterior equilibrium.[10] Moreover, any transfer scheme t, such that (x^*, t) permits implementation in posterior equilibrium, must have this form (Green and Laffont, 1977; Holmström, 1979).

Similarly, if preference types θ are fixed to be reported truthfully, and since agent i's expected utility given a profile of signals s is supermodular in (x, s_j) for each j, the sorting conditions of Proposition 4 of Bergemann and Välimäki (2002) are satisfied. Therefore, the transfer functions t such that (x^*, t) permits implementation in posterior equilibrium must belong to the class of generalized VCG mechanisms with transfers for agent 1 given by:[11]

$$\begin{aligned} \int_0^{s_1} \frac{\partial}{\partial x}\, \mathbb{E}_{\omega|v_1,s_2}\Big[u_2\big(x^*(\theta, v_1, s_2), \omega, \theta_2\big)\Big] \frac{\partial}{\partial v_1} x^*(\theta, v_1, s_2)\, dv_1 &= -(\theta_1 - \theta_2) \int_0^{s_1} \frac{\partial}{\partial v_1}\, \mathbb{E}_{\omega|v_1,s_2}[\omega]\, dv_1 \\ &= -(\theta_1 - \theta_2)\Big[\,\mathbb{E}_{\omega|s_1,s_2}[\omega] - \mathbb{E}_{\omega|0,s_2}[\omega]\Big]. \end{aligned}$$

Since $(\theta_1 - \theta_2)\,\mathbb{E}_{\omega|0,s_2}[\omega]$ does not depend on s_1, the generalized VCG payments for agent 1 are of the form

$$t_1(\theta, s; k_1) = k_1(s_2; \theta) - (\theta_1 - \theta_2)\,\mathbb{E}_{\omega|s}[\omega],$$

for an arbitrary function $k_1(s_2; \theta)$ of s_2.

Towards a contradiction, suppose there is a transfer rule t implementing the efficient outcome in posterior equilibrium. Since the transfer function t must ensure there are no profitable deviations along each dimension, it follows that there are functions h_1 and k_1 such that

$$h_1(\theta_2; s) - k_1(s_2; \theta) = \frac{1}{4}(\theta_1 - \theta_2)^2 + \mathrm{Var}_{\omega|s}[\omega] - (\theta_1 - \theta_2)\,\mathbb{E}_{\omega|s}[\omega].$$

Repeating the above for $s_1' \neq s_1$ and s_2, θ:

$$h_1(\theta_2; s_1', s_2) - k_1(s_2; \theta) = \frac{1}{4}(\theta_1 - \theta_2)^2 + \mathrm{Var}_{\omega|s_1',s_2}[\omega] - (\theta_1 - \theta_2)\,\mathbb{E}_{\omega|s_1',s_2}[\omega].$$

Taking their difference, we obtain

$$h_1(\theta_2; s) - h_1(\theta_2; s_1', s_2) = \mathrm{Var}_{\omega|s}[\omega] - \mathrm{Var}_{\omega|s_1',s_2}[\omega] - (\theta_1 - \theta_2)\Big(\,\mathbb{E}_{\omega|s}[\omega] - \mathbb{E}_{\omega|s_1',s_2}[\omega]\Big).$$

Since the right-hand side varies with θ_1 by $(*)$, while the left-hand side does not, we arrive at a contradiction, thus completing the proof.

[10] Equivalently, in the case signals are fixed and assumed not to be a strategic object, dominant strategies.

[11] This is up to the addition of an arbitrary function of agent 2's report of signal and the profile of preference types.

If a condition analogous to $(*)$ is not satisfied, implementation in posterior equilibrium remains possible. In a model with finite state and signal spaces, McLean and Postlewaite (2015) demonstrate that if agents have *zero informational size* in the terminology of McLean and Postlewaite (2002), that is, the private information held by any single agent is redundant when combined with the joint information of the other agents,[12] their generalized VCG mechanism allows for implementation in posterior equilibrium.[13] Moreover, McLean and Postlewaite (2015) show that if each agent exerts only a small informational effect on the posterior distribution over states, their generalized VCG mechanism is approximately posterior incentive-compatible.[14] In the limit, as the informational size of all agents approaches zero, the mechanism achieves exact implementation in posterior equilibrium.

Typically, however, we expect the signals of all agents to meaningfully influence predictions about the state. Our broader goal is to design mechanisms that are robust to assumptions on the underlying stochastic structure of signals and universally applicable across diverse settings. To this end, we now turn our attention to data-driven mechanisms.

4 Data-Driven Mechanisms

The previous section showed that there are instances where no message-driven mechanism can implement the efficient allocation in posterior equilibrium. In this section, we expand the class of mechanisms and demonstrate that this broader class enables at least approximate implementation, as formally defined below. To achieve this, we relax the assumption that both allocations and transfers depend solely on agents' reports, following an approach similar to Mezzetti (2004).[15] Instead, we assume that additional information about ω becomes available after the allocation is determined, allowing the designer to condition transfers on this information.

In practice, the designer acquires additional data about ω. For example, in the context of our motivating examples, this data might include information on user preferences, demand, click-through rates, or characteristics of an optimal output prompt in the context of LLMs. Such data can be collected by the platform based on users' interactions with the current environment, past and concurrent environments, or a combination thereof. This information might be derived from ad clicks, experiments designed to reveal user preferences, or demand estimation through revealed preference data sourced internally or externally. Moreover, in the context of

[12] McLean and Postlewaite (2015) refer to this as the *nonexclusive information* condition.

[13] In the generalized VCG (pivot) mechanism proposed by McLean and Postlewaite (2015), agent i pays the externality they impose on the other agents under the assumption that these agents have access to i's information even if i were absent. In the next section, we extend the standard VCG mechanism along similar lines. However, in our pivot mechanism, we account for two types of effects: first, the cost an agent imposes on others by influencing the allocation in their favor, and second, the value they contribute to improving prediction accuracy by sharing information about the state. For the latter, the agent is compensated, as illustrated in Example 2.

[14] The formalization of approximate posterior incentive compatibility by McLean and Postlewaite (2015) differs from our Definition 7. In their framework, a mechanism is defined as *weakly ϵ posterior incentive compatible* if truthful revelation is posterior incentive compatible up to an additive regret of ϵ, with a conditional probability of at least $1 - \epsilon$.

[15] See Subsection 4.2.2 for a comparison of Mezzetti (2004)'s approach with our framework.

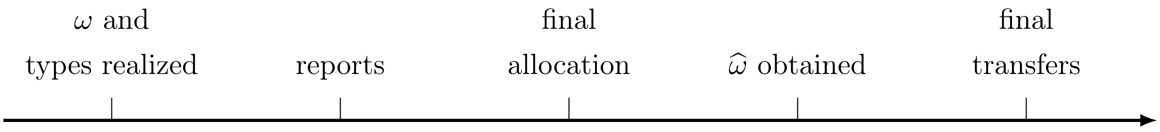

Figure 2: Timing of a data-driven direct revelation mechanism (Definition 4).

AI-generated content, user behaviors such as the time spent on specific prompts or follow-up queries could serve as valuable signals. While this data may not fully uncover user preferences, it still provides useful insights.

4.1 Mechanisms and Implementation

Let $\mathcal{D}$ represent the set of feasible datasets, transformations thereof, or any other form of additional information about ω. We define data-driven mechanisms as follows.

Definition 4. (Data-driven direct revelation mechanism)
A *data-driven direct revelation mechanism* is a pair (x, t) where $x : \Theta \times \mathcal{S} \to X$ is the outcome function and $t : \Theta \times \mathcal{S} \times \mathcal{D} \to \mathbb{R}^N$ are the transfers.

We formalize the availability of additional data by assuming the designer possesses an *estimator* of the state, with a realization obtained after the final allocation but before the final transfers are determined. With a slight abuse of notation, we denote by $\widehat{\omega}$ both the estimator and the estimate. The timeline is visualized in Figure 2.

Assumption 1. (Estimator)
The designer possesses an estimator $\widehat{\omega}$ of ω, with a realization obtained after the allocation but before final payments. The distribution of $\widehat{\omega} \mid \omega$ for any $\omega \in \Omega$ is common knowledge.

The definition of posterior equilibrium (Definition 3) extends naturally to the current setting by incorporating expected transfers, where the expectation is taken with respect to the additional information. Since agents do not observe the realization of the estimator at the reporting stage but know its data-generating process, they form beliefs about its realization.

Definition 5. (Posterior equilibrium with data-driven mechanisms)
A data-driven direct revelation mechanism (x, t) *permits implementation in posterior equilibrium* if for each agent $i \in N$ and true types $\theta \in \Theta$ and $s \in \mathcal{S}$:

$$v_i(x(\theta, s), \theta_i, s) + \bar{t}_i(\theta, s, s) \geq v_i(x(\theta_i', \theta_{-i}, s_i', s_{-i}), \theta_i, s) + \bar{t}_i(\theta_i', \theta_{-i}, s_i', s_{-i}, s), \forall s_i' \in \mathcal{S}_i, \theta_i' \in \Theta_i,$$

where $\bar{t}_i(s', \theta', s) \equiv \mathbb{E}_{\widehat{\omega}|s}[t_i(\theta', s', \widehat{\omega})]$.

With these definitions in place, we define a class of *data-driven VCG* mechanisms by allowing transfers to be conditioned on the realization of the estimator. This class of mechanisms will be the central focus of our analysis in this section.

Definition 6. (Data-driven VCG)
A data-driven direct revelation mechanism (x^*, t) is a *data-driven VCG mechanism* if for each i,

$$t_i(\theta, s, \widehat{\omega}) \equiv h_i(\theta_{-i}, s_{-i}, \widehat{\omega}) + \sum_{j \neq i} u_j(x^*(\theta, s), \widehat{\omega}, \theta_j),$$

for an arbitrary function h_i of others' reports and $\widehat{\omega}$ that is integrable with respect to $\widehat{\omega}$.

Any estimator $\widehat{\omega}$ defines a class of data-driven VCG mechanisms. This class is characterized by the specified transfer rule and the requirement that h_i be integrable with respect to the estimator's data-generating process for each agent i. Notably, regardless of the specific choice of the estimator, the class includes all transfer rules defined by functions h_i that are independent of the estimator for each agent i. Consequently, the intersection of these classes across all feasible estimators is non-empty.

Before continuing, we emphasize a central property of data-driven VCG transfers: they depend on the reported signals only through the resulting allocation. In other words, once the allocation is determined, the designer can compute the transfers without any knowledge of agents' beliefs.

We examine various desirable properties of the estimator and explore their implications for implementation in posterior equilibrium using data-driven VCG mechanisms. As a natural benchmark, we consider the ex-post case where $\widehat{\omega} = \omega$. If the designer ultimately observes the true state, ex-post utilities based on the reported preference types can be used to construct VCG payments, which are sufficient to achieve implementation in posterior equilibrium.

Proposition 1. *(Ex-post)*
If $\widehat{\omega} = \omega$, every data-driven VCG mechanism permits implementation in posterior equilibrium.

Proof. Fix an arbitrary data-driven VCG mechanism. At the reporting stage, each agent i evaluates the transfers in expectation. Toward achieving a posterior equilibrium at the reporting stage, assuming other agents report truthfully, for any θ, s and reports θ_i' and s_i', the expected transfer is given by

$$\begin{aligned} t_i(\theta_i', \theta_{-i}, s_i', s_{-i}) &\equiv \int_\Omega t_i(\theta_i', \theta_{-i}, s_i', s_{-i}, \omega)\, dP(\omega | s_i, s_{-i}) \\ &= h_i(\theta_{-i}, s_{-i}) + \sum_{j \neq i} v_j(x^*(\theta_i', \theta_{-i}, s_i', s_{-i}), \theta_j, s), \end{aligned}$$

where, with some abuse of notation, h_i is an arbitrary function of (θ_{-i}, s_{-i}). The net utility of agent i thus becomes

$$h_i(\theta_{-i}, s_{-i}) + \sum_{j=1}^{n} v_j(x^*(\theta_i', \theta_{-i}, s_i', s_{-i}), \theta_j, s). \tag{2}$$

By the efficiency of x^*, this expression is maximized at (θ_i, s_i). □

We illustrate the transfer scheme under the construction of Example 1.

Example 2. (Quadratic loss: data-driven VCG)
Example 1 continued. We also assume that the signals of the two agents are independent conditional on ω.

Consider the data-driven VCG transfers where we define

$$h_i(\theta_{-i}, s_{-i}, \omega) \equiv -u_j(x^*(\theta_j, s_j), \omega, \theta_j),$$

which corresponds to the pivot mechanism.

At the reporting stage, agent i evaluates the transfers in expectation. For any θ, s and reports θ_i' and s_i', the expected transfer is given by

$$\begin{aligned} t_i(\theta_i', \theta_j, s_i', s_j) &= v_j(x^*(\theta_i', \theta_j, s_i', s_j), \theta_j, s) - v_j(x^*(\theta_j, s_j), \theta_j, s) \\ &= \mathbb{E}_{\omega|s}\left[-\left(\frac{1}{2}\theta_i' - \frac{1}{2}\theta_j + \mathbb{E}_{\omega|s_i', s_j}[\omega] - \omega\right)^2 + \left(\mathbb{E}_{\omega|s_j}[\omega] - \omega\right)^2 \right]. \end{aligned}$$

As shown in Proposition 1, there is a posterior equilibrium where all agents report truthfully. In this case, the expected payment of agent i is

$$\begin{aligned} t_i(\theta, s) &= -\frac{1}{4}(\theta_i - \theta_j)^2 + \mathbb{E}_{\omega|s}\left[(\mathbb{E}_{\omega|s_j}[\omega] - \omega)^2 - (\mathbb{E}_{\omega|s}[\omega] - \omega)^2\right] \\ &= -\frac{1}{4}(\theta_i - \theta_j)^2 + \left(\mathbb{E}_{\omega|s}[\omega] - \mathbb{E}_{\omega|s_j}[\omega]\right)^2. \end{aligned}$$

We highlight the following properties of the payments.

Firstly, in expectation, the agent is *paid* for increasing the prediction accuracy. This is most clearly seen if preferences are aligned. Then the transfer of agent i is given by $(\mathbb{E}_{\omega|s_i,s_j}[\omega] - \mathbb{E}_{\omega|s_j}[\omega])^2$. This transfer is always non-negative and strictly positive if agent i's signal induces a different posterior mean when combined with agent j's signal than when based on agent j's signal alone. A necessary condition for this to hold is i having a strictly positive informational size in the terminology of McLean and Postlewaite (2002).

Secondly, the more accurate the information agent i provides, as measured by a reduction in the expected residual variance $\mathbb{E}[\text{Var}_{\omega|s_i,s_j}[\omega]|s_j] - \text{Var}_{\omega|s_j}[\omega]$, the higher is the agent's expected compensation. Intuitively, $(\mathbb{E}_{\omega|s_i,s_j}[\omega] - \mathbb{E}_{\omega|s_j}[\omega])^2$ measures how much additional information s_i provides about ω beyond s_j; the more informative the signal, the greater the deviation from the prior mean. Formally, this follows from the law of total variance:

$$\text{Var}\big[\,\mathbb{E}_{\omega|s_i,s_j}[\omega]|s_j\big] = \text{Var}_{\omega|s_j}[\omega] - \mathbb{E}\left[\text{Var}_{\omega|s_i,s_j}[\omega]|s_j\right],$$

noting that by the law of iterated expectations,

$$\text{Var}\big[\,\mathbb{E}_{\omega|s_i,s_j}[\omega]|s_j\big] = \mathbb{E}\left[\left(\,\mathbb{E}_{\omega|s}[\omega] - \mathbb{E}_{\omega|s_j}[\omega]\right)^2 |s_j\right].$$

Thus, a reduction in $\mathbb{E}[\text{Var}_{\omega|s_i,s_j}[\omega]|s_j] - \text{Var}_{\omega|s_j}[\omega]$ increases agent i's expected payment.

Thirdly, the agent *pays* for introducing bias into the social decision, with the expected payment for introducing bias given by $(\theta_i - \theta_j)^2/4$. When the bias is sufficiently large, the agent's total transfer becomes negative.

While the ex-post case offers a useful benchmark, it is more likely that the designer has only a noisy estimate of the state. We therefore consider estimators with two desirable properties commonly emphasized in statistics and econometrics: unbiasedness and consistency.

With an unbiased estimator and linearity of agents' utility functions in the state, data-driven VCG mechanisms achieve implementation in posterior equilibrium. The proof follows the same reasoning as in Proposition 1 and is omitted. By Jensen's inequality, this result does not extend universally beyond linear utilities.

Proposition 2. *(Unbiased estimator)*
If $\widehat{\omega}$ is an unbiased estimator of ω and the utility function of each agent is linear in ω, every data-driven VCG mechanism permits implementation in posterior equilibrium.

Next, we prove that data-driven VCG mechanisms with a consistent estimator of the state achieve implementation in ϵ-posterior equilibrium (formally defined below), where ϵ can be made arbitrarily small as the estimator converges in probability to the true state.

Definition 7. (ϵ-posterior equilibrium)
Fix $\epsilon \geq 0$. A data-driven direct revelation mechanism (x, t) *permits implementation in ϵ-posterior equilibrium* if for each agent $i \in N$ and true types $\theta \in \Theta$ and $s \in \mathcal{S}$:

$$v_i(x(\theta, s), \theta_i, s) + \bar{t}_i(\theta, s, s) + \epsilon \geq v_i(x(\theta_i', \theta_{-i}, s_i', s_{-i}), \theta_i, s) + \bar{t}_i(\theta_i', \theta_{-i}, s_i', s_{-i}, s), \forall s_i' \in \mathcal{S}_i, \theta_i' \in \Theta_i,$$

where $\bar{t}_i(s', \theta', s) \equiv \mathbb{E}_{\widehat{\omega}|s}[t_i(\theta', s', \widehat{\omega})]$.

In words, once the uncertainty about others' types is resolved, no agent regrets reporting truthfully by more than ϵ units of the numeraire.

Truthful reporting constitutes an ϵ-posterior equilibrium in finite samples for sufficiently large ϵ regardless of the estimator used. However, with a consistent estimator and under the smoothness conditions stated below, ϵ can be made arbitrarily small as the dataset size grows.[16]

Theorem 2. *(Consistent estimator)*
Suppose u_i is Lipschitz for each agent i and that the posterior $P(\cdot|s)$ is Lipschitz in the following sense:

$$\exists L_s > 0 \text{ such that } \forall s, s' \in \mathcal{S}, \delta_{TV}(P(\cdot|s), P(\cdot|s')) \leq L_s \|s - s'\|, \tag{3}$$

where δ_{TV} denotes the total variation distance. Consider a consistent estimator $\widehat{\omega}_m$ of ω, $\widehat{\omega}_m \to_p \omega$, conditional on any $\omega \in \Omega$. Then there is a sequence $\{\epsilon_m\}_m$ with $\epsilon_m \to 0$ as $m \to \infty$ such that every data-driven VCG mechanism for $\widehat{\omega}_m$ permits implementation in ϵ_m-posterior equilibrium for every $m \in \mathbb{N}$.

Proof. Fix an arbitrary $m \in \mathbb{N}$, a data-driven VCG mechanism for $\widehat{\omega}_m$, and true types $\theta \in \Theta$ and $s \in \mathcal{S}$. Let and $i \in N$ be arbitrary. Suppose agents other than i report truthfully.

[16] Note that our result also subsumes the following Bayesian interpretation. Suppose the platform observes signals $\rho_1, \ldots, \rho_m$ about ω. Updating the prior using Bayes rule, consider the posterior $\omega|\rho_1, \ldots, \rho_m$ and $\widehat{\omega}_m \equiv \mathbb{E}_{\omega|\rho_1,\ldots,\rho_m}[\omega]$ for each $m \in \mathbb{N}$. Under posterior consistency of the signals $\rho_1, \rho_2, \ldots, \rho_m, \ldots$, $\widehat{\omega}_m \to \omega$ almost surely. Hence, $\widehat{\omega}_m \to_p \omega$.

Using the law of iterated expectations, the expected transfer for agent i is

$$h_i(\theta_{-i}, s_{-i}) + \underbrace{\mathbb{E}_{\omega|s}\,\mathbb{E}_{\widehat{\omega}_m|\omega}\left[\sum_{j\neq i} u_j(x^*(\theta_i', \theta_{-i}, s_i', s_{-i}), \widehat{\omega}_m, \theta_j)\right]}_{\equiv t_i^m(x^*(\theta_i', \theta_{-i}, s_i', s_{-i}), \theta_{-i}, s)},$$

where, with some abuse of notation, h_i is an arbitrary function of (θ_{-i}, s_{-i}).

Let $\overline{x}^m \in X$ be a maximizer of this expression. By Weierstrass' theorem, this is well-defined due to continuity of the objective function and compactness. To simplify notation, define $\overline{x} \equiv x^*(\theta, s)$.

Define $\epsilon_i^m(\theta, s)$ as the utility loss from reporting truthfully:

$$\epsilon_i^m(\theta, s) \equiv v_i(\overline{x}^m, \theta_i, s) + t_i^m(\overline{x}^m, \theta_{-i}, s) - v_i(\overline{x}, \theta_i, s) - t_i^m(\overline{x}, \theta_{-i}, s) \geq 0.$$

Let $\epsilon_i^m \equiv \sup_{\theta,s} \epsilon_i^m(\theta, s)$, and $\epsilon_m \equiv \max_{i\in N} \epsilon_i^m$. By construction, truthful reporting constitutes an ϵ_m-posterior equilibrium. It remains to show that $\epsilon_m \to 0$ as $m \to \infty$.

Observe that $\epsilon_i^m(\theta, s)$ can be expressed as

$$\begin{aligned}\epsilon_i^m(\theta, s) = &\sum_{j=1}^n v_j(\overline{x}^m, \theta_j, s) + t_i^m(\overline{x}^m, \theta_{-i}, s) - \sum_{j\neq i} v_j(\overline{x}^m, \theta_j, s)\\ &- \sum_{j=1}^n v_j(\overline{x}, \theta_j, s) - t_i^m(\overline{x}, \theta_{-i}, s) + \sum_{j\neq i} v_j(\overline{x}, \theta_j, s).\end{aligned}$$

Moreover, by the efficiency of x,

$$\sum_{j=1}^n v_j(\overline{x}^m, \theta_j, s) \leq \sum_{j=1}^n v_j(\overline{x}, \theta_j, s).$$

Therefore,

$$\begin{aligned}\epsilon_i^m &= \sup_{\theta,s} \epsilon_i^m(\theta, s)\\ &\leq \sup_{\theta,s} t_i^m(\overline{x}^m, \theta_{-i}, s) - \sum_{j\neq i} v_j(\overline{x}^m, \theta_j, s) - t_i^m(\overline{x}, \theta_{-i}, s) + \sum_{j\neq i} v_j(\overline{x}, \theta_j, s).\end{aligned}$$

We obtain an upper bound:

$$\epsilon_i^m \leq 2 \sup_{\theta,s,x} \left| \sum_{j\neq i} v_j(x, \theta_j, s) - t_i^m(x, \theta_{-i}, s) \right|. \tag{4}$$

That $\sup_{\theta,s,x} \left| \sum_{j\neq i} v_j(x, \theta_j, s) - t_i^m(x, \theta_{-i}, s) \right|$ converges to 0 follows by Lemma 1 in Appendix A. Hence, ϵ_i^m converges to 0. Since i was chosen arbitrarily, it follows that ϵ_m also converges to 0 as $m \to \infty$. □

Next, we show that under stronger uniform integrability conditions, the sequence $\{\epsilon_m\}_m$ can converge to zero at essentially the same rate as the sequence of estimators converges to the true state: if $r_m(\widehat{\omega}_m - \omega)$ converges to zero in probability for a non-negative sequence $\{r_m\}_m$, $r_m\epsilon_m$ also converges to zero as $m \to \infty$. Up to a constant factor, this result provides an upper bound on the regret agents experience from reporting truthfully when others do the same, given a convergence rate of the estimator.

Proposition 3. *(Rate of convergence)*
Suppose u_i is Lipschitz for each agent i and that the posterior $P(\cdot|s)$ satisfies condition (3)*. Consider a sequence of estimators $\{\widehat{\omega}_m\}_m$ such that:*[17]

$$\widehat{\omega}_m - \omega = o_p\left(\frac{1}{r_m}\right).$$

Moreover, suppose $\{r_m(\widehat{\omega}_m - \omega)\}_m$ is a uniformly integrable sequence conditional on any $\omega \in \Omega$ and $\sup_{\omega\in\Omega}\sup_{m\in\mathbb{N}} \mathbb{E}_{\widehat{\omega}_m|\omega}[\|r_m(\widehat{\omega}_m - \omega)\|] < \infty$. Then there is a sequence of non-negative real numbers $\{\epsilon_m\}_m$ with:[18]

$$\epsilon_m = o\left(\frac{1}{r_m}\right),$$

such that every data-driven VCG mechanism for $\widehat{\omega}_m$ permits implementation in ϵ_m-posterior equilibrium for every $m \in \mathbb{N}$.

Proof. The proof is in Appendix A. □

4.2 Extensions and Discussion

4.2.1 Eliciting Additional Data from the Agents

Up to this point, we have assumed that the designer acquires additional information about the state through user engagement, feedback, or third-party sources. Alternatively, consider a different scenario: after the allocation is determined, each agent collects additional data or signals about the state, and a second reporting stage is introduced to elicit this information. Based on the reported data, we construct a *leave-one-out* estimator for each agent i, denoted by $\widehat{\omega}_{-i}$, which is derived solely from the information reported by the other agents. The corresponding data-driven VCG transfer for agent i is

$$t_i(\theta, s, \widehat{\omega}_{-i}) = h_i(\theta_{-i}, s_{-i}, \widehat{\omega}_{-i}) + \sum_{j\neq i} u_j(x^*(\theta, s), \widehat{\omega}_{-i}, \theta_j). \tag{5}$$

Analogously to the two-stage mechanism of Mezzetti (2004), truthful reporting in this second stage is a weakly dominant strategy for each agent: since an agent's transfer is independent of their own report, they are indifferent to their reporting choice. Working backward, given truthful reporting in the second stage, the first stage inherits the incentive properties of the induced games analyzed in this section, provided the corresponding leave-one-out estimators satisfy the required properties.[19] We state this as a result for the ex-post case.

[17]For a sequence of random variables X_n, we say $X_n = o_p(a_n)$ if for any $\epsilon > 0$, $\lim_{n\to\infty} \mathbb{P}\left(\frac{|X_n|}{a_n} > \epsilon\right) = 0$.

[18]For a sequence of non-negative numbers x_n, we say $x_n = o(a_n)$ if for every $\epsilon > 0$, there is $N \in \mathbb{N}$, such that $x_n \leq \epsilon a_n$ for all $n \geq N$.

[19]There are alternative formulations of the adapted data-driven VCG transfers. For instance, it is not necessary to aggregate the reported information from the second stage into a single leave-one-out estimator for each agent. Instead, the mechanism could elicit a separate estimate of the state from each agent and use the estimate reported by agent j in constructing the payments for all other agents. The key property remains that agent i's transfer does not depend on i's own reported estimate. Nonetheless, by combining information from multiple agents, the designer can accelerate the convergence of the corresponding estimator, enhancing its accuracy.

Corollary 1. *(Two-stage data-driven VCG in the ex-post case)*
Suppose each agent observes the true state after the final allocation but before the final transfers are determined. Consider a two-stage data-driven VCG mechanism: in the first stage, agents report their types, and in the second stage, they report their information about the state. Transfers are as in (5)*. Under this setup, truthfully reporting their type in the first stage and their information about the state in the second stage constitutes a Perfect Bayesian Equilibrium.*

4.2.2 Alternative Estimators and Information Requirements

We formalized the revelation of additional information about the state through an estimator of the state. An alternative approach would be to base data-driven transfers on estimators of *agent utilities*. With "perfect data," a similar result is obtained if we assume that agents' final utilities, $u_j(x^*(\theta', s'), \omega, \theta_j)$, are observed for each j instead of ω, using

$$t_i(\theta', s', u) \equiv h_i(\theta'_{-i}, s'_{-i}, u_{-i}) + \sum_{j \neq i} u_j(x^*(\theta', s'), \omega, \theta_j).$$

Mezzetti (2004) assumes agents observe their payoffs and report them in a second stage. With reports u'_j for all j:

$$t_i(\theta', s', u') \equiv h_i(\theta'_{-i}, s'_{-i}, u'_{-i}) + \sum_{j \neq i} u'_j.$$

Reporting truthfully at both stages constitutes a Perfect Bayesian Equilibrium.

Our data-driven mechanisms require *strictly less information* for implementation. Specifically, the final payments do not depend on the knowledge of the agents' true preference types to align incentives. To illustrate, consider the transfer rules in the ex-post case. The "core" of our data-driven VCG transfers for each agent i, based on the reports (θ', s'), is given by:

$$\sum_{j \neq i} u_j(x^*(\theta', s'), \omega, \theta'_j),$$

whereas the Mezzetti (2004)-style transfers rely on the knowledge of the agents' *true* payoffs for the chosen allocation. This distinction is further highlighted by comparing our Corollary 1 with Proposition 1 of Mezzetti (2004), which explicitly requires such payoff information.

From a practical standpoint, our data-driven mechanisms also show implementation in posterior equilibrium is possible in cases where a second reporting stage is impractical or prohibitively costly. In such situations, we argue that it is more natural to use the approach developed in this paper. Estimators of agents' payoffs could be derived in the following scenario: the mechanism is executed, the platform implements the chosen allocation, and the user interacts with the output. Based on user engagement and the relative "prominence" of each agent in the output, the platform can observe the user's interaction with individual agents or at least obtain a noisy estimate, potentially supplemented by off-platform data.[20] Using this information, the platform can estimate each agent's resulting payoff. However, this interpretation relies on the assumption that the platform has an estimate of user engagement—the unobserved component of agents' payoffs. Therefore, it seems more natural to analyze estimators of the state rather than utilities.

[20] For example, the platform may not directly observe users' purchasing behavior after clicking an ad, but could access external indicators of such actions.

4.2.3 State-Revealing Signals

The results in this and the previous section apply to scenarios where agents' signals, when combined, fully reveal the state (see Footnote 8 for a discussion in the context of Theorem 1). This implies that the designer cannot rely solely on reported signals and preference types to determine transfers. However, when an estimator with the properties discussed in this section is available, the designer can condition transfers on it to align incentives. Importantly, even if the estimator is noisy, it cannot simply be replaced by a state constructed directly from agents' reports, as this would allow scope for manipulating transfers through reported signals.

4.2.4 Multiple Payments

All results presented in this section remain valid if we split the payment into multiple stages, provided that the total payment matches the corresponding transfer schemes. For instance, at the reporting stage, the designer could require a payment equal to the expected VCG payment based on the reported preference type. Once $\widehat{\omega}$ is observed, the final transfers would then be adjusted to match the corresponding data-driven transfers.

4.2.5 Interdependent Preferences

The arguments presented in this section no longer hold if we introduce additional interdependence into agents' payoffs. Specifically, if the payoff function u_i of each agent i depends on the entire type profile $\theta \in \Theta$, such that

$$u_i : X \times \Theta \times \Omega \to \mathbb{R},$$

the proofs no longer go through. Under these conditions, agent i's report of θ_i influences the expected data-driven VCG transfer not only indirectly, through its effect on the allocation, but also directly, by entering others' payoff functions. Consequently, the main results of this section no longer hold. This failure is illustrated in Example 5 in Appendix A.

5 Applications

5.1 Click-Through Auctions

In this section, we apply our framework to a canonical model of click-through auctions (Edelman et al., 2007; Varian, 2007). To keep the analysis simple, we consider an auction for a single advertising slot. The space of feasible allocations is the probability simplex $X = \Delta_{N-1}$. The state is given by the *click-through rates* (CTR) for each agent conditional on winning the slot: $\Omega = [0, 1]^N$ with a typical element $\omega = (\omega_1, \dots, \omega_N)$.[21] Each agent i's preference type θ_i is defined as i's value per click, i.e., $\Theta_i \subseteq \mathbb{R}_+$. Agents are assumed to be risk-neutral, with payoffs given by:

$$u_i(x, \omega, \theta_i) = \theta_i \cdot x_i \cdot \omega_i.$$

[21] Alternatively, we could define ω_i to be the expected number of clicks without substantially changing the analysis.

We assume a stationary environment, where clicks on displayed ads are modeled as independent draws from a Bernoulli distribution with parameters equal to the true click-through probabilities.

Recall that in message-driven direct revelation mechanisms, the allocation and transfers depend only on the reported types. This class includes per-impression payments but excludes per-click payments. By following steps similar to Example 1 and assuming the strictly positive informational size condition ($*$), we can show that no message-driven mechanism implements the efficient allocation in posterior equilibrium in all instances.

Per-click payments belong to the broader class of mechanisms that employ data-driven transfers, with clicks serving as proxies for user engagement and revealing valuable information. Under the maintained assumptions, the observed sample click-through rate becomes an unbiased and consistent estimator of the true click-through rate.

Implementation with per-click pivot (second-price) payments depends on whether the click-through rates are common or individual. First, consider the case of a common click-through rate, where $\omega = \omega_i = \omega_j$ for each $i, j \in N$. In an efficient allocation, the agent with the highest value per click wins the slot; in the case of ties, any tie-breaker rule is efficient. Consider the per-click pivot payment t^{pc} for agent i: $t_i^{pc}(\theta, s) = \max_{j \neq i} \theta_j$.

Corollary 2. *(Common CTR pivot payments)*
Suppose there is a common click-through rate. Then the mechanism (x^*, t^{pc}) with the per-click pivot payment t^{pc} permits implementation in posterior equilibrium.

Proof. We show that the expected per-click pivot and data-driven pivot transfers with an estimator given by the observed click-through rate coincide. Therefore, they must provide the same incentives at the reporting stage. The result will then follow from Proposition 2.

Fix an arbitrary agent i and reports (θ', s'). Agent i expects the ad to be clicked with probability $\mathbb{E}_{\omega|s}[\omega]$ conditional on being allocated the slot, which happens with probability $x_i(\theta', s')$. Thus, the expected per-click transfer is

$$\max_{j \neq i} \theta'_j \cdot x_i^*(\theta', s') \cdot \mathbb{E}_{\omega|s}[\omega].$$

Now consider the data-driven pivot transfers:

$$\max_{j \neq i} \theta'_j \cdot x_i^*(\theta', s') \cdot \mathbb{E}_{\widehat{\omega}|s}[\widehat{\omega}].$$

By the law of iterated expectations and the fact that the estimator is unbiased

$$\mathbb{E}_{\widehat{\omega}|s}[\widehat{\omega}] = \mathbb{E}_{\omega|s} \mathbb{E}_{\widehat{\omega}|\omega}[\widehat{\omega}] = \mathbb{E}_{\omega|s}[\omega].$$

Therefore, the expected data-driven pivot payment is

$$\max_{j \neq i} \theta'_j \cdot x_i^*(\theta', s') \cdot \mathbb{E}_{\omega|s}[\omega].$$

The result follows. □

Next, consider the case of individual click-through rates. The efficient allocation assigns the slot to agent i if i's expected payoff is the largest among agents, i.e.,

$$\theta_i \cdot \mathbb{E}_{\omega_i|s}[\omega_i] > \max_{j \neq i} \left\{ \theta_j \cdot \mathbb{E}_{\omega_j|s}[\omega_j] \right\}.$$

In case of ties, any tie-breaking rule is efficient.

However, in this case, the per-click pivot payment does not align incentives. Indeed, fix an arbitrary agent i and assume the other agents report truthfully. Under the per-click pivot payment, i's expected payment would be

$$\max_{j \neq i} \theta_j \cdot x_i^*(\theta_i', \theta_{-i}, s_i', s_{-i}) \cdot \mathbb{E}_{\omega_i|s}[\omega_i].$$

Observe that the expected payment depends on the reported signals solely through the allocation. Consequently, agent i may have an incentive to misreport signals, inflating her own expected click-through rate while diminishing those of others. By doing so, agent i can secure greater prominence by promising higher payment frequencies, even though, based on her posterior beliefs, these payments are unlikely to materialize. We illustrate with the following example.

Example 3. (Individual CTR pivot payments)
Consider two agents, and suppose that while the prior on ω_i has non-zero variance, each agent knows her own click-through rate, $\omega_i = s_i$, for each i. Assume further that $\theta_1 s_1 < \theta_2 s_2$, but $\theta_1 > \theta_2$. If agent 2 reports truthfully and agent 1 reports (θ_1', s_1'), the expected net utility of agent 1 under the per-click pivot payment is given by

$$(\theta_1 - \theta_2) x_1^*(\theta_1', \theta_2, s_1', s_2) s_1 = (\theta_1 - \theta_2) \mathbb{1}_{\{\theta_1' s_1' \geq \theta_2 s_2\}} s_1.$$

It follows that agent 1 optimally chooses to report $s_1', \theta_1' = 1$. Truthful reporting by all agents is not a posterior equilibrium.

With individual click-through rates, the corresponding expected data-driven pivot payment for agent i would take the form

$$t_i(\theta', s') = \max_{j \neq i} \left\{ \theta_j' \cdot \mathbb{E}_{\omega_j|s}[\omega_j] \right\} \cdot x_i^*(\theta', s'). \tag{6}$$

However, implementing such transfers as per-click payments based on agent i's clicks is infeasible, as it would require knowledge of the ratios ω_j/ω_i for each $j \in N \setminus \{i\}$ or an unbiased or consistent estimator of them. If agent i is awarded the slot, the designer would not have a consistent or unbiased estimator of other agents' click-through rates from the user interaction with the current environment, as other ads are not viewed. Nevertheless, the data-driven pivot mechanism can be implemented if the platform has an estimator of ω_j, potentially sourced from other auction environments or third-party sources.[22]

[22] We could construct payments with expectations matching (6) based on clicks in the current auction environment, though possibly at the expense of reducing efficiency. For instance, consider two agents and a lottery that selects the efficient allocation with probability $p \in (0, 1)$, assigning the slot randomly otherwise. Suppose the efficient allocation gives the slot to agent i. Under the lottery, agent i receives the slot with probability $\frac{1+p}{2}$, and agent j with probability $\frac{1-p}{2}$. Setting i's payment as

$$\frac{2}{1-p} \theta_j, \tag{7}$$

whenever agent j's ad is clicked when assigned the slot (occurring with probability ω_j), makes i's expected payment match (6). This method implements the inefficient allocation with probability $\frac{1-p}{2}$, which can be made arbitrarily close to 0 as p approaches 1. However, note that as p approaches 1, the payment conditional on clicks in (7) diverges to infinity.

5.2 Mechanism Design for LLM-Generated Content

We now apply our framework and results to a mechanism design setting where content is generated by an LLM. LLMs operate as autoregressive generation systems, sequentially predicting the conditional probability of the next token based on the token sequence generated thus far (Brown, 2020).[23] These probabilities are modeled by neural networks trained to maximize the likelihood of the observed text in the training data. For the purpose of our mechanism design framework, we remain agnostic about the specific unit of text output by the LLM—whether it is an individual token or a sequence of tokens forming an output prompt. We collectively refer to this as the "output text". An intrinsic aspect of our analysis is that output texts are randomly generated.

In this context, each input prompt defines a unique mechanism design environment. The output of the LLM is modeled as a unit of text that responds to this prompt. Let T denote the set of feasible outputs, with a typical element $t \in T$. We assume this set is finite. Each agent $i \in N$ is endowed with a *reward function* r_i, which depends on the generated output text t and the *relevant context* or query-specific information $\omega \in \Omega$. The context captures elements such as user intent or a hidden "ground truth" associated with the prompt. The reward functions are assumed to be be continuous and common knowledge up to a finite-dimensional parameter $\theta_i \in \Theta_i$, which represents the agent's private *preference type*. Each agent i also possesses a *private signal* $s_i \in \mathcal{S}_i$, which improves the accuracy and relevance of the LLM's output. Signals represent agents' training data or knowledge bases in the context of the Retrieval-Augmented Generation framework (Lewis et al., 2020). Such datasets may overlap, and information across agents may be correlated, as advertisers may share segments of their customer base or source data from the same providers. An important aspect of our model is that it allows for arbitrary correlation among agents' signals. The sets $\Omega, \Theta, \mathcal{S}$ follow our earlier specification.

The set of outcomes X corresponds to the set of *generation distributions over output text units*, denoted by $\Delta(T)$. For a given input prompt, architecture, rewards, and data, an LLM produces a generation distribution $x \in X$. Define the *ex-post utility* of agent i from an LLM's generation distribution x as $u_i : X \times \Omega \times \Theta_i \to \mathbb{R}$ such that

$$u_i(x, \omega, \theta_i) = \sum_{t \in T} r_i(t, \omega, \theta_i) \cdot x(t).$$

This specification aligns directly with the general framework introduced earlier. We define v_i based on u_i as in Section 2.

Finally, we extend the framework to include a *reference LLM*, $x_0(s)$, representing the platform's organic generation distribution, similarly to Soumalias et al. (2024). We let the reference generation distribution condition on agents' data. The reference LLM's objective is to provide responses aligned with user queries and maximize their usefulness. However, the platform permits deviations from the reference LLM's output if the agents seeking to steer the output according to their preferences make sufficiently high payments. These payments serve as a mech-

[23] Tokens, the fundamental units of LLM-generated text, can include subwords, words, phrases, symbols, or numbers. LLMs generate text incrementally on a *token-by-token* basis.

anism to balance the competing objectives of maintaining useful, high-quality responses while accommodating agents' influence, distorting the output.

We assume the designer aims to determine a central LLM generation distribution that maximizes the sum of agents' payoffs while penalizing deviations from the reference LLM distribution.[24] An analogous objective was previously used by Soumalias et al. (2024) and follows standard approaches for tuning LLMs towards human preferences, such as the Reinforcement Learning from Human Feedback (RLHF) (Ouyang et al., 2022) and Direct Preference Optimization (DPO) (Rafailov et al., 2024).[25]

Definition 8. (α-regularized efficient generation distribution)
The LLM generation distribution $x : \Theta \times \mathcal{S} \to X$ is α-regularized efficient if, for all $\theta \in \Theta$ and $s \in \mathcal{S}$, it satisfies

$$x(\theta, s) \in \arg\max_{x \in X} \mathbb{E}_{\omega|s} \left[\sum_{t \in T} \sum_{i \in N} r_i(t, \omega, \theta_i) x(t) \right] - \alpha \rho\left(x, x_0(s)\right) \tag{8}$$

$$= \arg\max_{x \in X} \sum_{i \in N} v_i(x, \theta_i, s) - \alpha \rho\left(x, x_0(s)\right). \tag{9}$$

where $\rho : X \times X \to \mathbb{R}_+$ measures the deviation of x from x_0.

Fixing α, we denote by x^* an α-regularized efficient generation distribution.

Following the RLHF and DPO literature (Ouyang et al., 2022; Rafailov et al., 2024), Soumalias et al. (2024) specify ρ as the Kullback-Leibler divergence $D_{KL}(x \| x_0) = \sum_{t \in T} x(t) \log \frac{x(t)}{x_0(t)}$. We keep the framework general and return to this functional form in Example 4 below. We assume throughout that ρ is continuous in x. Note that by continuity and compactness, the α-regularized efficient generation distribution is always well-defined.

The designer's goal is to implement an α-regularized efficient generation distribution in posterior equilibrium. We assume the platform collects feedback from user engagement with the central LLM's output prompt and leverages it to construct an estimator of ω. Given an estimate $\widehat{\omega}$ of ω, the modified data-driven VCG transfers are defined as follows.

Definition 9. (α-regularized data-driven VCG)
A data-driven direct revelation mechanism (x^*, t) is an *α-regularized data-driven VCG mecha-*

[24] Another approach would be to specify the platform as an additional agent with a reward function r_0. This function, dependent on t and ω, is assumed to be common knowledge. The results below can be readily adapted to incorporate this framework. Specifically, rather than introducing a regularization term for deviations from the reference distribution in the social objective function, we could redefine the social objective as a weighted sum of agents' reward functions—including the reward function of the reference LLM agent. This redefinition would remain subject to standard constraints imposed by the LLM architecture. The resulting objective would then guide the specification of a generation distribution that maximizes the constrained social welfare. The data-driven VCG mechanism could be adjusted accordingly by appropriately weighting the payoffs of agents other than the focal agent. We decided to use the model presented in this section as it aligns closely with the discussed related literature in computer science.

[25] Unlike in RLHF and DPO, we optimize the sum of many agents' rewards under incentive conflicts, whereas the focus of these papers is optimizing a single agent's reward without considering incentives.

nism if for each i:

$$t_i(\theta, s, \widehat{\omega}) \equiv h_i(\theta_{-i}, s_{-i}, \widehat{\omega}) + \sum_{j \neq i} u_j(x^*(\theta, s), \widehat{\omega}, \theta_j) - \alpha\rho\left(x^*(\theta, s), x_0(s)\right).$$

for an arbitrary function h_i of others' reports and $\widehat{\omega}$ that is integrable with respect to $\widehat{\omega}$.

Under appropriate regularity conditions on the reward functions, all results from Section 4 can be readily extended to the current setting. This is formalized in the following corollary, where the statements correspond to Proposition 1, Proposition 2, Theorem 2, and Proposition 3, respectively. For brevity, we omit the proof.

Corollary 3. *(Implementation with α-regularized data-driven VCG)*
The following holds:

1. *If $\widehat{\omega} = \omega$, every α-regularized data-driven VCG mechanism permits implementation in posterior equilibrium.*

2. *If $\widehat{\omega}$ is an unbiased estimator of ω and the reward function of each agent is linear in ω, every α-regularized data-driven VCG mechanism permits implementation in posterior equilibrium.*

3. *Suppose r_i is Lipschitz for each agent i and that the posterior $P(\cdot|s)$ satisfies condition* (3). *Consider a consistent estimator $\widehat{\omega}_m$ of ω, $\widehat{\omega}_m \to_p \omega$, conditional on any $\omega \in \Omega$. Then there is a sequence $\{\epsilon_m\}_m$ with $\epsilon_m \to 0$ as $m \to \infty$ such that every α-regularized data-driven VCG mechanism for $\widehat{\omega}_m$ permits implementation in ϵ_m-posterior equilibrium for every $m \in \mathbb{N}$.*

4. *If, in addition to the conditions in 2., $\widehat{\omega}_m - \omega = o_p\left(\frac{1}{q_m}\right)$, $\{q_m(\widehat{\omega}_m - \omega)\}_m$ is a uniformly integrable sequence conditional on any $\omega \in \Omega$, and $\sup_{\omega\in\Omega} \sup_{m\in\mathbb{N}} \mathbb{E}_{\widehat{\omega}_m|\omega}[\|q_m(\widehat{\omega}_m - \omega)\|] < \infty$, the statement of part 2. holds with $\epsilon_m = o\left(\frac{1}{q_m}\right)$.*

We illustrate the construction using D_{KL} as the regularization function.

Example 4. (α-regularized data-driven VCG with KL regularization)
Let $\rho = D_{KL}$. Following the derivation by Peters and Schaal (2007) and Rafailov et al. (2024), among others, the α-regularized efficient generation distribution is given by

$$x^*(t|\theta, s) = \frac{x_0(t|s)}{Z(\theta, s)} \exp\left(\frac{1}{\alpha}\sum_{i\in N} \mathbb{E}_{\omega|s}[r_i(t, \omega, \theta_i)]\right),$$

for each $t \in T$, where

$$Z(\theta, s) \equiv \sum_{t\in T} x_0(t|s) \exp\left(\frac{1}{\alpha}\sum_{i\in N} \mathbb{E}_{\omega|s}[r_i(t, \omega, \theta_i)]\right)$$

is the partition function. As derived in Appendix A, the maximized social objective in (8) is $\alpha \log\left(Z(\theta, s)\right)$ and the α-regularized data-driven pivot transfer of agent i is

$$t_i(\theta, s, \widehat{\omega}) = \alpha[\log\left(Z(\theta, s)\right) - \log\left(Z(\theta_{-i}, s_{-i}\right)] - v_i(x^*(\theta, s), \theta_i, s)$$
$$+\sum_{j\neq i} [u_j(x^*(\theta, s), \widehat{\omega}, \theta_j) - v_j(x^*(\theta, s), \theta_j, s)] - [u_j(x^*(\theta_{-i}, s_{-i}), \widehat{\omega}, \theta_j) - v_j(x^*(\theta_{-i}, s_{-i}), \theta_j, s_{-i})].$$

Note that in the ex-post case, where $\widehat{\omega} = \omega$, and assuming a posterior equilibrium in which all agents report truthfully, the expected transfer to agent i simplifies to

$$\begin{aligned} t_i(\theta, s) = &\sum_{j \neq i} v_j(x^*(\theta, s), \theta_j, s) - v_j(x^*(\theta_{-i}, s_{-i}), \theta_j, s_{-i}) \\ &+ \alpha\Big(D_{KL}(x^*(\theta_{-i}, s_{-i}) \| x_0(s_{-i})) - D_{KL}(x^*(\theta, s) \| x_0(s))\Big) \\ = &\, \alpha[\log(Z(\theta, s)) - \log(Z(\theta_{-i}, s_{-i})] - v_i(x^*(\theta, s), \theta_i, s). \end{aligned}$$

That is, agent i's expected net payoff is the agent's marginal contribution to the social objective in (8). The intuition from Example 2 extends to this setting as well. Informally, if agents' preference types are sufficiently aligned and agent i's information is valuable for the joint prediction problem, we expect that, ex-ante with respect to signal realizations, the expression

$$\sum_{j \neq i} v_j(x^*(\theta, s), \theta_j, s) - v_j(x^*(\theta_{-i}, s_{-i}), \theta_j, s_{-i})$$

is positive in expectation. Further, if the difference of the regularization terms

$$D_{KL}(x^*(\theta, s) \| x_0(s)) - D_{KL}(x^*(\theta_{-i}, s_{-i}) \| x_0(s_{-i}))$$

is relatively small compared to the welfare gain for the other agents, that is, if agent i does not introduce substantial marginal bias into the output of the central LLM, agent i's total transfer is positive.[26] We also note that the agent has to pay for introducing bias through two channels: (i) by imposing an allocation externality on the other agents, and (ii) by steering the generation distribution away from the reference LLM's distribution.

6 Discussion and Conclusion

We offered an approach to mechanism design that harnesses the natural flow of information in digital environments. By conditioning transfers on post-allocation data, we showed how to achieve implementation even in challenging multidimensional settings. Our framework provides a foundation for designing practical mechanisms in modern applications where rich feedback data is readily available.

We also highlighted connections to the literature on efficient mechanism design with interdependent values. In particular, with respect to Mezzetti (2004)'s results, our results provide an insight into the kind and amount of post-allocation data necessary to ensure implementation in at least approximate posterior equilibrium.

This work opens several promising directions for future research, two of which we highlight here. First, although we focused on implementation in posterior equilibrium, exploring implementation in Bayesian equilibrium would be valuable. We anticipate that a similar impossibility result for message-driven mechanisms would hold with independent types, following the results

[26] For instance, when the reward functions of all agents are perfectly aligned with the reward function of the reference LLM, the α-regularized efficient generation distribution coincides with the reference LLM's generation distribution. Consequently, the difference between the two KL divergence terms becomes zero.

of Maskin (1992), Dasgupta and Maskin (2000), and Jehiel and Moldovanu (2001). However, there may be potential for implementing the efficient allocation with correlated types, using constructions similar to the full-surplus-extraction mechanisms of Crémer and McLean (1988) and McAfee and Reny (1992). These could be combined with a version of VCG payments if at least one component of the types satisfies Crémer and McLean (1988)'s conditions on the stochastic structure. In such cases, consistent with Neeman (2004)'s insight, agents would retain some information rents. Yet, when these stochastic relevance conditions are not met, it remains unclear whether standard message-driven mechanisms can effectively align incentives.

Second, we assumed a fixed information structure for agents. From a practical perspective, incentivizing agents to acquire the socially optimal amount of costly information is an important desideratum, as the central platform aggregates decentralized information from individual agents. Modeling costly information acquisition as an additional stage in the induced game and designing a mechanism that incentivizes the efficient level of information acquisition is an important direction for future research.

References

Bergemann, D., Duetting, P., Paes Leme, R., and Zuo, S. (2022). Calibrated click-through auctions. In *Proceedings of the ACM Web Conference 2022*, pages 47–57.

Bergemann, D. and Välimäki, J. (2002). Information acquisition and efficient mechanism design. *Econometrica*, 70(3):1007–1033.

Braverman, M. and Chassang, S. (2022). Data-driven incentive alignment in capitation schemes. *Journal of Public Economics*, 207:104584.

Brown, T. B. (2020). Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*.

Caragiannis, I., Kaklamanis, C., Kanellopoulos, P., Kyropoulou, M., Lucier, B., Leme, R. P., and Tardos, É. (2015). Bounding the inefficiency of outcomes in generalized second price auctions. *J. Econ. Theory*, 156:343–388.

Chen, J., Li, M., Xu, H., and Zuo, S. (2023). Bayesian calibrated click-through auction. *arXiv preprint arXiv:2306.06554*.

Clarke, E. H. (1971). Multipart pricing of public goods. *Public choice*, pages 17–33.

Crémer, J. and McLean, R. P. (1988). Full extraction of the surplus in bayesian and dominant strategy auctions. *Econometrica: Journal of the Econometric Society*, pages 1247–1257.

Dasgupta, P. and Maskin, E. (2000). Efficient auctions. *The Quarterly Journal of Economics*, 115(2):341–388.

Dubey, K. A., Feng, Z., Kidambi, R., Mehta, A., and Wang, D. (2024). Auctions with llm summaries. *arXiv preprint arXiv:2404.08126*.

Dütting, P., Fischer, F. A., and Parkes, D. C. (2019). Expressiveness and robustness of first-price position auctions. *Math. Oper. Res.*, 44(1):196–211.

Dütting, P., Fischer, F. A., and Parkes, D. C. (2024). Nontruthful position auctions are more robust to misspecification. *Math. Oper. Res.*, 49(2):901–927.

Dütting, P., Mirrokni, V., Paes Leme, R., Xu, H., and Zuo, S. (2024). Mechanism design for large language models. In *Proceedings of the ACM on Web Conference 2024*, pages 144–155.

Edelman, B., Ostrovsky, M., and Schwarz, M. (2007). Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American economic review*, 97(1):242–259.

Green, J. and Laffont, J.-J. (1977). Characterization of satisfactory mechanisms for the revelation of preferences for public goods. *Econometrica: Journal of the Econometric Society*, pages 427–438.

Green, J. R. and Laffont, J.-J. (1987). Posterior implementability in a two-person decision problem. *Econometrica: Journal of the Econometric Society*, pages 69–94.

Groves, T. (1973). Incentives in teams. *Econometrica: Journal of the Econometric Society*, pages 617–631.

Hajiaghayi, M., Lahaie, S., Rezaei, K., and Shin, S. (2024). Ad auctions for llms via retrieval augmented generation. *arXiv preprint arXiv:2406.09459*.

Hansen, R. G. (1985). Auctions with contingent payments. *The American Economic Review*, 75(4):862–865.

Holmström, B. (1979). Groves' scheme on restricted domains. *Econometrica: Journal of the Econometric Society*, pages 1137–1144.

Jehiel, P., Meyer-ter Vehn, M., Moldovanu, B., and Zame, W. R. (2007). Posterior implementation vs ex-post implementation. *Economics Letters*, 97(1):70–73.

Jehiel, P. and Moldovanu, B. (2001). Efficient design with interdependent valuations. *Econometrica*, 69(5):1237–1259.

Jehiel, P. and Moldovanu, B. (2005). Allocative and informational externalities in auctions and related mechanisms.

Klemperer, P. (1998). Auctions with almost common values: Thewallet game'and its applications. *European Economic Review*, 42(3-5):757–769.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., et al. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, 33:9459–9474.

Liang, A. and Madsen, E. (2024). Data and incentives. *Theoretical Economics*, 19(1):407–448.

Lu, J. (2019). Bayesian identification: a theory for state-dependent utilities. *American Economic Review*, 109(9):3192–3228.

Maskin, E. (1992). Auctions and privatization. *Privatization*, 1:15–136.

McAfee, R. P. and Reny, P. J. (1992). Correlated information and mecanism design. *Econometrica: Journal of the Econometric Society*, pages 395–421.

McLean, R. and Postlewaite, A. (2002). Informational size and incentive compatibility. *Econometrica*, 70(6):2421–2453.

McLean, R. P. and Postlewaite, A. (2015). Implementation with interdependent valuations. *Theoretical Economics*, 10(3):923–952.

McLean, R. P. and Postlewaite, A. (2017). A dynamic non-direct implementation mechanism for interdependent value problems. *Games and Economic Behavior*, 101:34–48.

Mezzetti, C. (2004). Mechanism design with interdependent valuations: Efficiency. *Econometrica*, 72(5):1617–1626.

Milgrom, P. (2010). Simplified mechanisms with an application to sponsored-search auctions. *Games and Economic Behavior*, 70(1):62–70.

Milgrom, P. R. and Weber, R. J. (1982). A theory of auctions and competitive bidding. *Econometrica: Journal of the Econometric Society*, pages 1089–1122.

Neeman, Z. (2004). The relevance of private information in mechanism design. *Journal of Economic theory*, 117(1):55–77.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., et al. (2022). Training language models to follow instructions with human feedback. *Advances in neural information processing systems*, 35:27730–27744.

Peters, J. and Schaal, S. (2007). Reinforcement learning by reward-weighted regression for operational space control. In *Proceedings of the 24th international conference on Machine learning*, pages 745–750.

Rafailov, R., Sharma, A., Mitchell, E., Manning, C. D., Ermon, S., and Finn, C. (2024). Direct preference optimization: Your language model is secretly a reward model. *Advances in Neural Information Processing Systems*, 36.

Riordan, M. H. and Sappington, D. E. (1988). Optimal contracts with public ex post information. *Journal of Economic Theory*, 45(1):189–199.

She, Z., Ayer, T., and Montanera, D. (2022). Can big data cure risk selection in healthcare capitation program? a game theoretical analysis. *Manufacturing & Service Operations Management*, 24(6):3117–3134.

Soumalias, E., Curry, M. J., and Seuken, S. (2024). Truthful aggregation of llms with an application to online advertising. *arXiv preprint arXiv:2405.05905*.

Varian, H. R. (2007). Position auctions. *international Journal of industrial Organization*, 25(6):1163–1178.

Varian, H. R. (2009). Online ad auctions. *American Economic Review*, 99(2):430–434.

Varian, H. R. and Harris, C. (2014). The vcg auction in theory and practice. *American Economic Review*, 104(5):442–445.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1):8–37.

Wu, J., Badanidiyuru, A., and Xu, H. (2024). Auctioning with strategically reticent bidders. In *Web and Internet Economics: 20th International Conference, WINE 2024*. Springer.

A Proofs and Further Results

Lemma 1. *Suppose u_i is Lipschitz for each agent i and that the posterior $P(\cdot|s)$ is Lipschitz in the following sense: $\exists L_s > 0$ such that for any $s, s' \in \mathcal{S}$:*

$$\delta_{TV}(P(\cdot|s), P(\cdot|s')) \leq L_s \|s - s'\|,$$

where δ_{TV} denotes the total variation distance. Consider a consistent estimator $\widehat{\omega}_m$ of ω conditional on ω, i.e. $\widehat{\omega}_m \to_p \omega$. Then for each $i \in N$:

$$\lim_{m\to\infty} \sup_{\theta_i\in\Theta_i, x\in X, s\in\mathcal{S}} \left| v_i(x, \theta_i, s) - \int_\omega \mathbb{E}_{\widehat{\omega}_m|\omega}\Big[u_i(x, \widehat{\omega}_m, \theta_i)\Big] dP(\omega|s) \right| = 0.$$

Proof. Fix arbitrary $i \in N$. First, note that since u_i is continuous, by the continuous mapping theorem, $\forall x \in X$ and $\theta_i \in \Theta_i$, $u_i(x, \widehat{\omega}_m, \theta_i) \to_p u_i(x, \omega, \theta_i)$. Moreover, the sequence of $u_i(x, \widehat{\omega}_m, \theta_i)$ is uniformly integrable for any $x \in X, \theta_i \in \Theta_i$, which follows from continuity and compactness of Ω.[27] Hence, for any $\omega \in \Omega$:

$$\mathbb{E}_{\widehat{\omega}_m|\omega}\left[u_i(x, \widehat{\omega}_m, \theta_i)\right] \xrightarrow{L_1} u_i(x, \omega, \theta_i).$$

By the Bounded Convergence Theorem, for any $\theta_i \in \Theta_i, x \in X, s \in \mathcal{S}$:

$$\lim_{m\to\infty} \int_\Omega \mathbb{E}_{\widehat{\omega}_m|\omega}\left[u_i(x, \widehat{\omega}_m, \theta_i)\right] dP(\omega|s) = \int_\Omega u_i(x, \omega, \theta_i) dP(\omega|s) = v_i(x, \theta_i, s). \tag{11}$$

We claim the convergence is uniform. To this end, we prove the following claims.

Fix an arbitrary $m \in \mathbb{N}$. Define the mapping $\Psi_m : X \times \Theta_i \times \mathcal{S} \to \mathbb{R}$ by

$$(x, \theta_i, s) \mapsto \int_\Omega \mathbb{E}_{\widehat{\omega}_m|\omega}\Big[u_i(x, \widehat{\omega}_m, \theta_i)\Big] dP(\omega|s).$$

The family of functions $\{\Psi_m\}_m$ is uniformly equicontinuous.[28] Indeed, let $\epsilon > 0$ be arbitrary and fix $\delta = \frac{\epsilon}{\sqrt{2}(ML_s + L_u)}$, where $M = \|u_i\|_\infty$, L_s is the Lipschitz constant on the posterior with respect to s and L_u is the Lipschitz constant on u_i.

[27] A sequence of random variables $\{X_n\}_n$ is uniformly integrable if

$$\lim_{\lambda\to\infty} \limsup_{n\to\infty} \mathbb{E}[|X_n| \mathbb{1}_{|X_n|\geq\lambda}] = 0. \tag{10}$$

[28] A family of functions $\mathcal{F}$ between metric spaces (X, d_x) and (Y, d_y) is uniformly equicontinuous if $\forall\epsilon > 0, \exists\delta > 0$ such that $\forall f \in \mathcal{F}$ and $x_1, x_2 \in X$ with $d_x(x_1, x_2) < \delta$, $d_x(f(x_1), f(x_2)) < \epsilon$.

Observe that for any $m \in \mathbb{N}$,

$$\begin{aligned}
&|\Psi_m(x_1,\theta_{i1},s_1) - \Psi_m(x_2,\theta_{i2},s_2)| \\
&= \left| \int_\Omega \mathbb{E}_{\widehat{\omega}_m|\omega}\left[u_i(x_1,\widehat{\omega}_m,\theta_{i1})\right] dP(\omega|s_1) - \int_\Omega \mathbb{E}_{\widehat{\omega}_m|\omega}\left[u_i(x_2,\widehat{\omega}_m,\theta_{i2})\right] dP(\omega|s_2)\right| \\
&= \left| \int_\Omega \mathbb{E}_{\widehat{\omega}_m|\omega}\left[u_i(x_1,\widehat{\omega}_m,\theta_{i1})\right] dP(\omega|s_1) - \int_\Omega \mathbb{E}_{\widehat{\omega}_m|\omega}\left[u_i(x_1,\widehat{\omega}_m,\theta_{i1})\right] dP(\omega|s_2)\right. \\
&\left. + \int_\Omega \mathbb{E}_{\widehat{\omega}_m|\omega}\left[u_i(x_1,\widehat{\omega}_m,\theta_{i1})\right] dP(\omega|s_2) - \int_\Omega \mathbb{E}_{\widehat{\omega}_m|\omega}\left[u_i(x_2,\widehat{\omega}_m,\theta_{i2})\right] dP(\omega|s_2)\right| \\
&\leq \left| \int_\Omega \mathbb{E}_{\widehat{\omega}_m|\omega}\left[u_i(x_1,\widehat{\omega}_m,\theta_{i1})\right]\left[dP(\omega|s_1) - dP(\omega|s_2)\right]\right| \\
&+ \left| \int_\Omega \mathbb{E}_{\widehat{\omega}_m|\omega}\left[u_i(x_1,\widehat{\omega}_m,\theta_{i1}) - u_i(x_2,\widehat{\omega}_m,\theta_{i2})\right] dP(\omega|s_2)\right| \\
&\equiv I + II.
\end{aligned}$$

To bound I, note that by out continuity and compactness assumptions, and by properties of the total variation distance,

$$I \leq \|u_i\|_\infty \delta_{TV}(P(\cdot|s_1) - P(\cdot|s_2)) \leq ML_s\|s_1 - s_2\|.$$

To bound II, note that, since $|\cdot|$ is convex, by Jensen's inequality:

$$II \leq \int_\Omega \mathbb{E}_{\widehat{\omega}_m|\omega}\left[\left|u_i(x_1,\widehat{\omega}_m,\theta_{i1}) - u_i(x_2,\widehat{\omega}_m,\theta_{i2})\right|\right] dP(\omega|s_2).$$

Further, by Lipschitz continuity of u_i and since probability measures integrate to 1:

$$II \leq L_u\|(x_1,\theta_{i1}) - (x_2,\theta_{i2})\|.$$

Combining the above,

$$\begin{aligned}
I + II &\leq ML_s\|s_1 - s_2\| + L_u\|(x_1,\theta_{i1}) - (x_2,\theta_{i2})\| \\
&\leq (ML_s + L_u)\left(\|s_1 - s_2\| + \|(x_1,\theta_{i1}) - (x_2,\theta_{i2})\|\right) \\
&\leq \sqrt{2}(ML_s + L_u)\|(x_1,\theta_{i1},s_1) - (x_2,\theta_{i2},s_2)\|,
\end{aligned}$$

which is smaller than ϵ by our assumptions. Moreover, this holds uniformly across m. Since $\epsilon > 0$ was arbitrary, it follows that $\{\Psi_m\}$ is uniformly equicontinuous (hence, for each $m \in \mathbb{N}$, Ψ_m is continuous).

Further, define the mapping $\Psi : X \times \Theta_i \times \mathcal{S} \to \mathbb{R}$ by $(x, \theta_i, s) \mapsto v_i(x, \theta_i, s)$. By an analogous argument to the preceding paragraph, Ψ is continuous.

Further, note that the sequence $\{\Psi_m\}_m$ is uniformly bounded by $\|u_i\|_\infty$.

As noted in (11), $\Psi_m \to \Psi$ pointwise. But since $X \times \Theta_i \times \mathcal{S}$ is compact, $\{\Psi_m\}$ and Ψ belong to the set of continuous functions on $X \times \Theta_i \times \mathcal{S}$, and the family of functions $\{\Psi_m\}_m$ is uniformly equicontinuous and uniformly bounded, it follows by the Arzela–Ascoli theorem that the convergence in (11) is uniform. Thus,

$$\lim_{m\to\infty} \sup_{\theta_i\in\Theta_i, x\in X, s\in\mathcal{S}} \left| v_i(x,\theta_i,s) - \int_\omega \mathbb{E}_{\widehat{\omega}_m|\omega}\left[u_i(x,\widehat{\omega}_m,\theta_i)\right] dP(\omega|s)\right| = 0.$$

Since $i \in N$ was arbitrary, the result follows. $\square$

Proof of Proposition 3

Proof. Fix an arbitrary data-driven VCG mechanism, agent i, and true $\theta \in \Theta$ and $s \in \mathcal{S}$.

Recall from (4) that for each $i \in N$ and $m \in \mathbb{N}$,

$$\epsilon_i^m \leq 2 \sup_{\theta_i, s, x} \left| \sum_{j \neq i} v_j(x) - t_i^m(x) \right| \leq 2 \sum_{j \neq i} \sup_{\theta_i, s, x} \left| \mathbb{E}_{\omega|s}[u_j(x, \omega, \theta_j)] - \mathbb{E}_{\widehat{\omega}_m|s}[u_j(x, \widehat{\omega}_m, \theta_j)] \right|.$$

For each $j \in N$, applying Jensen's inequality and the law of iterated expectations,

$$\begin{aligned} \left| \mathbb{E}_{\omega|s}[u_j(x, \omega, \theta_j)] - \mathbb{E}_{\widehat{\omega}_m|s}[u_j(x, \widehat{\omega}_m, \theta_j)] \right| &= \left| \mathbb{E}_{\omega|s} \left[u_j(x, \omega, \theta_j) - \mathbb{E}_{\widehat{\omega}_m|\omega}[u_j(x, \widehat{\omega}_m, \theta_j)] \right] \right| \\ &\leq \mathbb{E}_{\omega|s} \left[\left| u_j(x, \omega, \theta_j) - \mathbb{E}_{\widehat{\omega}_m|\omega}[u_j(x, \widehat{\omega}_m, \theta_j)] \right| \right] \\ &= \mathbb{E}_{\omega|s} \left[\left| \mathbb{E}_{\widehat{\omega}_m|\omega}[u_j(x, \omega, \theta_j) - u_j(x, \widehat{\omega}_m, \theta_j)] \right| \right] \\ &\leq \mathbb{E}_{\omega|s} \mathbb{E}_{\widehat{\omega}_m|\omega} \left[\left| u_j(x, \omega, \theta_j) - u_j(x, \widehat{\omega}_m, \theta_j) \right| \right]. \end{aligned}$$

By Lipschitz continuity of u_j, it follows that u_j is Lipschitz continuous in ω uniformly in x, θ_j: there is $L_j \geq 0$ such that for each $x \in X$ and $\theta_j \in \Theta_j$,

$$|u_j(x, \omega, \theta_j) - u_j(x, \widehat{\omega}_m, \theta_j)| \leq L_j \|\omega - \widehat{\omega}_m\|.$$

Hence,

$$\left| \mathbb{E}_{\omega|s}[u_j(x, \omega, \theta_j)] - \mathbb{E}_{\widehat{\omega}_m|s}[u_j(x, \widehat{\omega}_m, \theta_j)] \right| \leq L_j \, \mathbb{E}_{\omega|s} \mathbb{E}_{\widehat{\omega}_m|\omega} \left[\|\omega - \widehat{\omega}_m\| \right]$$

and

$$\sup_{\theta_j, x} \left| \mathbb{E}_{\omega|s}[u_j(x, \omega, \theta_j)] - \mathbb{E}_{\widehat{\omega}_m|s}[u_j(x, \widehat{\omega}_m, \theta_j)] \right| \leq L_j \, \mathbb{E}_{\omega|s} \mathbb{E}_{\widehat{\omega}_m|\omega} \left[\|\omega - \widehat{\omega}_m\| \right].$$

Finally,

$$\epsilon_i^m \leq 2 \sup_s \sum_{j \neq i} L_j \, \mathbb{E} \left[\|\omega - \widehat{\omega}_m\| | s \right],$$

and so

$$r_m \epsilon_i^m \leq 2 \sup_s \sum_{j \neq i} L_j \, \mathbb{E} \left[\|r_m(\omega - \widehat{\omega}_m)\| | s \right].$$

Since $r_m(\omega - \widehat{\omega}_m) = o_p(1)$, $\{r_m(\omega - \widehat{\omega}_m)\}_m$ is a uniformly integrable sequence conditional on any ω, $\sup_{\omega \in \Omega} \max_{m \in \mathbb{N}} \mathbb{E}_{\widehat{\omega}_m|\omega}[\|r_m(\widehat{\omega}_m - \omega)\|] < \infty$, it follows by the Bounded convergence theorem that

$$\lim_{m \to \infty} \mathbb{E} \left[\|r_m(\omega - \widehat{\omega}_m)\| | s \right] = 0,$$

for any $s \in \mathcal{S}$. It remains to be show that this convergence is uniform. The proof proceeds analogously to the proof of Theorem 2. In particular, it can be shown that the family of mappings $s \mapsto \mathbb{E} \left[\|r_m(\omega - \widehat{\omega}_m)\| | s \right]$ for all $m \in \mathbb{N}$ is uniformly equicontinuous, using that the integrand is uniformly bounded and that the posterior is Lipschitz continuous in signals. We omit the details for brevity.

It follows that $\epsilon_i^m = o\left(\frac{1}{r_m}\right)$ for every $i \in N$. Therefore, $\epsilon_m = \max_{i \in N} \epsilon_i^m = o\left(\frac{1}{r_m}\right)$, concluding the proof. $\square$

Example 5. Suppose there are two agents and a single object to be allocated. The set of allocations is the probability simplex Δ_1. The state, preference types, and signals are single-dimensional and all belonging to $[0, 1]$. Suppose there are no atoms in the priors on types.

Suppose the payoff of agent 1 is given by

$$u_1(x, \theta, \omega) = \theta_1 \cdot x_1 \cdot \omega,$$

where x_1 is the probability agent 1 obtains the object. The payoff of agent 2 is given by

$$u_2(x, \theta, \omega) = (-\theta_1 + \theta_2) \cdot x_2 \cdot \omega.$$

The efficient allocation is as follows. It is efficient to allocate the object to agent 1 if $2\theta_1 > \theta_2$. If $2\theta_1 < \theta_2$, it is efficient to allocate the object to agent 2. In the case of a tie, any feasible allocation is efficient.

Consider the data-driven VCG expected transfer for the ex-post case:

$$t_1(\theta, s) = h_1(\theta_2, s_2) + \mathbb{E}_{\omega|s}\left[u_2(x^*(\theta, s), \theta, \omega)\right].$$

Suppose agent 2 reports truthfully. Agent 1's expected net utility when reporting (θ_1', s_1') thus becomes

$$h_1(\theta_2, s_2) + \left[\theta_1 x_1^*(\theta_1', \theta_2) + (-\theta_1' + \theta_2) x_2^*(\theta_1', \theta_2)\right] \mathbb{E}_{\omega|s}\left[\omega\right].$$

If $\theta_1 < \theta_2$ and $\mathbb{E}_{\omega|s}[\omega] > 0$, it is optimal for agent 1 to report $\theta_1' = 0$, yielding a payoff of $h_1(\theta_2, s_2) + \theta_2 \mathbb{E}_{\omega|s}[\omega]$ Therefore, reporting truthfully for all agents is not a posterior equilibrium.

Derivations for Example 4

Recall the maximizer of (8) is

$$x^*(t|\theta, s) = \frac{x_0(t|s)}{Z(\theta, s)} \exp\left(\frac{1}{\alpha} \sum_{i \in N} \mathbb{E}_{\omega|s}[r_i(t, \omega, \theta_i)]\right).$$

To compress notation, define

$$R(t|\theta, s) \equiv \exp\left(\frac{1}{\alpha} \sum_{i \in N} \mathbb{E}_{\omega|s}[r_i(t, \omega, \theta_i)]\right).$$

Plugging in $x^*(\theta, s)$ into the social objective in 8 yields

$$\begin{aligned}
&\sum_{t \in T} \sum_{i \in N} \mathbb{E}_{\omega|s}\left[r_i(t, \omega, \theta_i)\right] \frac{x_0(t|s)}{Z(\theta, s)} R(t|\theta, s) - \alpha \sum_{t \in T} \frac{x_0(t|s)}{Z(\theta, s)} R(t|\theta, s) \log\left(\frac{R(t|\theta, s)}{Z(\theta, s)}\right) \\
&= \sum_{t \in T} \sum_{i \in N} \mathbb{E}_{\omega|s}\left[r_i(t, \omega, \theta_i)\right] \frac{x_0(t|s)}{Z(\theta, s)} R(t|\theta, s) - \sum_{t \in T} \sum_{i \in N} \mathbb{E}_{\omega|s}\left[r_i(t, \omega, \theta_i)\right] \frac{x_0(t|s)}{Z(\theta, s)} R(t|\theta, s) \\
&\quad + \alpha \log\left(Z(\theta, s)\right) \sum_{t \in T} \frac{x_0(t|s)}{Z(\theta, s)} R(t|\theta, s) \\
&= \alpha \log\left(Z(\theta, s)\right).
\end{aligned}$$

Now consider

$$\sum_{j\neq i} u_j(x^*(\theta,s),\widehat{\omega},\theta_j) - \alpha D_{KL}(x^*(\theta,s)\|x_0(s)),$$

where, as just derived,

$$\begin{aligned}\alpha D_{KL}(x^*(\theta,s)\|x_0(s)) &= \sum_{t\in T}\sum_{i\in N}\mathbb{E}_{\omega|s}\left[r_i(t,\omega,\theta_i)\right]\frac{x_0(t|s)}{Z(\theta,s)}R(t|\theta,s) - \alpha\log\left(Z(\theta,s)\right)\\ &= \sum_{i\in N} v_i(x^*(\theta,s),\theta_i,s) - \alpha\log\left(Z(\theta,s)\right).\end{aligned}$$

Also note that

$$\sum_{j\neq i} u_j(x^*(\theta,s),\widehat{\omega},\theta_j) = \sum_{t\in T}\sum_{j\neq i} r_j(t,\widehat{\omega},\theta_j)\frac{x_0(t|s)}{Z(\theta,s)}R(t|\theta,s).$$

Thus,

$$\begin{aligned}&\sum_{j\neq i} u_j(x^*(\theta,s),\widehat{\omega},\theta_j) - \alpha D_{KL}(x^*(\theta,s)\|x_0(s))\\ &= \alpha\log\left(Z(\theta,s)\right) - v_i(x^*(\theta,s),\theta_i,s) + \sum_{j\neq i} u_j(x^*(\theta,s),\widehat{\omega},\theta_j) - v_j(x^*(\theta,s),\theta_j,s).\end{aligned}$$

Similarly,

$$\begin{aligned}&\sum_{j\neq i} u_j(x^*(\theta_{-i},s_{-i}),\widehat{\omega},\theta_j) - \alpha D_{KL}(x^*(\theta_{-i},s_{-i})\|x_0(s_{-i}))\\ &= \alpha\log\left(Z(\theta_{-i},s_{-i})\right) + \sum_{j\neq i} u_j(x^*(\theta_{-i},s_{-i}),\widehat{\omega},\theta_j) - v_j(x^*(\theta_{-i},s_{-i}),\theta_j,s_{-i}).\end{aligned}$$

Therefore, the α-regularized data-driven pivot transfer of agent i is

$$\begin{aligned}&\alpha[\log\left(Z(\theta,s)\right) - \log\left(Z(\theta_{-i},s_{-i}\right)] - v_i(x^*(\theta,s),\theta_i,s)\\ +&\sum_{j\neq i}\left[u_j(x^*(\theta,s),\widehat{\omega},\theta_j) - v_j(x^*(\theta,s),\theta_j,s)\right] - \left[u_j(x^*(\theta_{-i},s_{-i}),\widehat{\omega},\theta_j) - v_j(x^*(\theta_{-i},s_{-i}),\theta_j,s_{-i})\right].\end{aligned}$$