Persuasive Prediction via Decision Calibration

Jingwu Tang*
Carnegie Mellon University
jingwutang@cmu.edu

Jiahao Zhang*
Carnegie Mellon University
jiahaozhang@cmu.edu

Fei Fang
Carnegie Mellon University
feifang@cmu.edu

Steven Wu
Carnegie Mellon University
zstevenwu@cmu.edu

Abstract

Bayesian persuasion, a central model in information design, studies how a sender, who privately observes a state drawn from a prior distribution, strategically sends a signal to influence a receiver's action. A key assumption is that both sender and receiver share the precise knowledge of the prior. Although this prior can be estimated from past data, such assumptions break down in high-dimensional or infinite state spaces, where learning an accurate prior may require a prohibitive amount of data. In this paper, we study a learning-based variant of persuasion, which we term *persuasive prediction*. This setting mirrors Bayesian persuasion with large state spaces, but crucially does not assume a common prior: the sender observes covariates X, learns to predict a payoff-relevant outcome Y from past data, and releases a prediction to influence a population of receivers. To model rational receiver behavior without a common prior, we adopt a learnable proxy: *decision calibration*, which requires the prediction to be unbiased conditioned on the receiver's best response to the prediction. This condition guarantees that myopically responding to the prediction yields no swap regret. Assuming the receivers best respond to decision-calibrated predictors, we design a provably efficient algorithm that learns a decision-calibrated predictor within a randomized predictor class that optimizes the sender's utility. In the commonly studied single-receiver case, our method matches the utility of a Bayesian sender who has full knowledge of the underlying prior distribution and is restricted to signaling schemes induced by the same predictor class. Finally, we extend our algorithmic result to a setting where receivers respond stochastically to predictions and the sender may randomize over an infinite predictor class.

1 Introduction

The strategic disclosure of information to influence downstream decisions—commonly studied under the *information design* literature (see e.g., Bergemann and Morris [2019])—has become a central topic in economic theory. A foundational model in this area is *Bayesian persuasion*, introduced by Kamenica and Gentzkow [2011], which formalizes how a *sender*, who privately observes an underlying state unknown to the *receiver*, can send a signal to shape the receiver's belief and ultimately influence their chosen action. Crucially, the sender is endowed with the *power of commitment*—the ability to credibly commit to a signaling scheme in advance. Based on the committed signaling scheme, the receiver updates their beliefs upon receiving the signal and selects an action accordingly.

*Equal contribution.

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

Both receivers have utility functions that depend on the state and the receiver's action, but these utilities are typically misaligned, so that the sender needs to strategically design their scheme.

A key assumption underpinning Bayesian persuasion—and much of information design—is the existence of a *common prior*: a probability distribution over the state space that both the sender and receiver have precise knowledge about. The most obvious source for such a probability distribution is past data. However, in many real-world environments, the state space may be high-dimensional or even infinite, which makes it challenging to estimate an accurate prior distribution from finite data. A natural example is the following setting of prediction for decision making.

Persuasive prediction problem. The data $(x, y) \in \mathcal{X} \times \mathcal{Y}$ are drawn from a joint distribution $\mathcal{D}$. The sender observes a realization x of $\mathcal{X}$, and seeks to predict the unobserved outcome y. Both the sender's and the receiver's utilities depend on the receiver's action and the unknown outcome Y, but not directly on X. Upon observing x, the sender provides a prediction $f(x)$, which acts as a signal to inform the receiver's decision. In the language of Bayesian persuasion, the sender's observed state corresponds to the posterior distribution $\mathbb{P}[Y \mid X = x]$, and the prior is the marginal distribution over such posterior distributions, which is effectively $\mathcal{D}$. However, when $\mathcal{X}$ is high-dimensional or infinite, learning the prior $\mathcal{D}$ from finite past data can be infeasible.

This persuasive prediction framework captures a wide range of real-world scenarios where the sender observes high-dimensional covariates describing the state of the world and aims to shape individual decisions toward a socially desirable outcome. For example, a small community bank may partner with a fintech platform that outputs predicted default probabilities based on high-dimensional financial and behavioral data. However, the incentives of the two parties are not fully aligned: the fintech platform benefits from increasing loan origination volume and thus may have incentives to produce predictions that make approvals more likely, whereas the bank remains cautious and primarily uses these predictions to safeguard against excessive default risk.

When data are insufficient to recover the full prior, persuasive prediction presents two central challenges: *How should receiver behavior be modeled in the absence of a common prior? And can the sender, using only finite data, achieve utility comparable to that of a fully informed Bayesian sender with exact knowledge of the distribution $\mathcal{D}$?*

1.1 Our results and techniques

In persuasive prediction, a sender learns a predictor f from past data such that upon observing $X = x$, they will send the prediction $f(x)$ to a population of receivers, who will then select their actions. The sender's goal is to optimize their utility that depends jointly on the unknown outcome Y and the joint action chosen by the receivers. Our results establish a connection between (Bayesian) persuasion and *decision calibration* [Noarov et al., 2023, Zhao et al., 2021], which allows a learning-based approach for modeling incentives without requiring full knowledge of the prior.

Behavioral modeling via decision calibration. Informally, a (possibly randomized) predictor $f\colon \mathcal{X} \to \mathcal{Y}$ is *decision-calibrated* if, for every receiver with utility function v_i and every action a_i,

$$\mathbb{E}_{f,\mathcal{D}}\big[\, Y - f(X) \;\big|\; \arg\max_a v_i(f(X), a) = a_i \,\big] = 0. \tag{1}$$

Intuitively, decision calibration captures a natural notion of credibility: conditioned on any event defined by the receiver's best response, the predictor $f(X)$ must be an unbiased estimate of the true outcome Y. A predictor is approximately decision-calibrated if condition equation 1 holds up to a small additive error. We show that myopically best responding to an approximately decision-calibrated, receivers obtain low swap regret (Theorem 2.5). This motivates a clean behavioral assumption that receivers best respond to approximately decision calibrated predictions.

Efficient optimal persuasive prediction. Suppose the sender is allowed to use a stochastic predictor $f \in \Delta(\mathcal{H})$, that is, a distribution over deterministic predictors in some finite class $\mathcal{H}$. Our first main result is a statistically efficient algorithm that learns a predictor f that optimizes the sender's utility within the class of decision-calibrated predictors in $\Delta(\mathcal{H})$. The core technical idea is to formulate the learning problem as a zero-sum game between a *min player*, who updates the predictor, and an *max player*, who identifies the most violated calibration constraint. Simulating no-regret dynamics between the two players yields a minimax equilibrium, which recovers the optimal decision-calibrated

predictor. The number of required samples scales polynomially in the number of receiver actions and the dimension of $\mathcal{Y}$, and is independent of the size $|\mathcal{X}|$. Our algorithm is also *oracle-efficient* in the sense that it runs in polynomial time when given access to an ERM oracle over $\mathcal{H}$.

Matching the Bayesian benchmark. In the special case of a single receiver, we show that the sender utility achieved by our algorithm matches that of a fully informed Bayesian sender who is restricted to sending signals induced by the same class of decision-calibrated predictors, even though our sender only has access to a finite dataset that is far from sufficient to approximate $\mathcal{D}$.

Extension to quantal responses. Finally, we extend our results to settings where receivers perform *quantal responses*-—that is, their action choices follow a softmax distribution rather than deterministic best responses. This extension models scenarios where receiver behavior is stochastic and not perfectly rational [McKelvey and Palfrey, 1995]. Under this setting, we also provide an efficient algorithm for learning the approximately optimal decision-calibrated predictor for sender utility, and can also handle infinite hypothesis classes $\mathcal{H}$, provided it has bounded covering numbers. We empirically validate this extension on a real-world lending dataset, considering both single-receiver and multi-receiver settings; due to space constraints, we defer the details to Appendix H.

1.2 Related work

The work most closely related to ours is Feng and Tang [2025], which studies the problem of selecting an optimal calibrated predictor to maximize the learner's utility. Their formulation builds on the connection between online calibration and Bayesian persuasion established by Jain and Perchet [2024]. However, the settings and objectives of the two works differ in important ways. Feng and Tang consider a *full-information* optimization problem in which the sender knows the underlying distribution $\mathcal{D}$ and computes an optimal predictor subject to standard calibration constraints. In contrast, we study a prior-free learning problem in which $\mathcal{D}$ is unknown, motivated by settings where the context space $\mathcal{X}$ is sufficiently rich that learning the full conditional distribution $D_{\mathcal{Y}|x}$ from finite data is infeasible. We also differ in the calibration notion considered. While Feng and Tang [2025] focuses on standard calibration, we optimize over the larger class of decision-calibrated predictors. Because decision calibration is formally weaker, it can potentially achieve higher sender utility. At the same time, unlike a standard calibrated predictor, a decision-calibrated predictor does not directly induce a Bayesian persuasion signal scheme; relating the two requires collapsing predictive values within each best-response region, as formalized in Lemma 4.3.

Due to space limitations, we include additional related work in the Appendix A.

2 Model and preliminaries

Predictors We consider the prediction task over the data domain $\mathcal{X} \times \mathcal{Y}$, where the data is drawn from a distribution $\mathcal{D}$. Here, $\mathcal{X}$ is a rich feature space, and $\mathcal{Y} = [-1, 1]^d$ is the outcome space. We define $\mathcal{H} = \{h \mid h : \mathcal{X} \to \mathcal{Y}\}$ as a hypothesis class of *deterministic* predictors. For any $h \in \mathcal{H}$ and $x \in \mathcal{X}$, $h(x)$ is interpreted as a prediction of the conditional mean $\mathbb{E}[Y \mid X = x]$. We use $h(x)_j$ and y_j to denote the j-coordinate of the predicted and true outcome vectors, respectively.

In this paper, we consider a more general setting where the goal is to learn a *randomized* predictor $f \in \Delta(\mathcal{H})$, representing a distribution over $h \in \mathcal{H}$. This aligns with the standard information design literature, where a sender typically transmits a randomized signal to the receiver to achieve higher utility. We assume no direct access to the full distribution $\mathcal{D}$; instead, we seek to learn f from n i.i.d. samples drawn from $\mathcal{D}$, which we denote as dataset D.

In our model, the receivers observe only the released prediction $f(x)$ and their own utility function, but not x at decision time. This is analogous to standard Bayesian persuasion, where the receiver does not observe the realized state at decision time and therefore best-responds to the sender's signal.

Receivers' behavior model We consider N receivers who make decisions based on the prediction $h(x)$. For each $i \in [N]$, receiver i has a finite action set $\mathcal{A}_i$. Without loss of generality, we assume $|\mathcal{A}_i| = m$ for all $i \in [N]$, since any smaller set can be augmented with dummy actions to reach size m. Receiver i's utility function is denoted as $v_i(a, y)$, where $v_i \colon \mathcal{A}_i \times \mathcal{Y} \longrightarrow [0, 1]$.

We assume that v_i is linear and Lipschitz continuous in the outcome y. Note that this mild because in the classical Bayesian persuasion setting with a finite state space, the sender's expected utility is naturally linear in the posterior mean.

Assumption 2.1 (Linearity and L-Lipschitzness)**.** *For any $i \in [N]$, and $a \in \mathcal{A}_i$, the utility function $v_i(a, y)$ is linear in y, and satisfies $|v_i(a, y_1) - v_i(a, y_2)| \leq L\|y_1 - y_2\|_\infty$.*

Next, we define the receiver's decision rule given the prediction $h(x)$. A natural rule is to treat the prediction as accurate and respond optimally to it.

Definition 2.2 (Strict Best Response)**.** *For any $i \in [N]$, receiver i, given utility function v_i, strictly best responds to the prediction $h(x)$ by choosing:*

$$b_i(h(x), a_i) = \begin{cases} 1 & \text{if } a_i = \arg\max_{a_i' \in \mathcal{A}_i} v_i(a_i', h(x)), \\ 0 & \text{otherwise.} \end{cases}$$

Here, $b_i(h(x), a_i)$ represents the probability that receiver i takes action a_i given the prediction $h(x)$.

Decision calibration We aim to design f such that receivers experience no regret when best responding to it. This mirrors the setting in standard Bayesian persuasion, where, given a known prior, the sender recommends an action through a signal, and the receiver's best response—after Bayesian updating—is to follow that recommendation. As we do not assume a known prior, we leverage the notion of *decision calibration* [Noarov et al., 2023, Zhao et al., 2021], which has been shown to provide similar no-regret guarantees for receivers. Specifically, it ensures that receivers have no incentive to deviate from the recommended action, whether by swapping actions or by acting as if their utility were that of another receiver. Moreover, it can be shown that any approximately decision-calibrated predictor can be post-processed into an approximately fully calibrated predictor that induces the same receiver's behavior. We will show that fully calibrated predictors correspond to signaling schemes in Bayesian persuasion with a known prior (Lemma 4.2). A similar observation was made in Jain and Perchet [2024] for online calibrated predictions. But they made finite-state Bayesian persuasion assumption, which we do not require. For these reasons, we adopt decision calibration as a desirable property that the predictor f should satisfy. We provide a detailed discussion of the no-regret guarantees in Appendix C, and their connection to Bayesian persuasion in Section 4.

We now formally define decision calibration as follows.

Definition 2.3 (Decision Calibration)**.** *For a randomized predictor $f \in \Delta(\mathcal{H})$, define*

$$g_{i,j,a_i}(h) := \mathbb{E}_{(x,y)\sim\mathcal{D}}\big[(y_j - h(x)_j)\, b_i(h(x), a_i)\big].$$

We say that f is perfectly decision calibrated *if*

$$\mathrm{DecCE}(f) := \max_{i\in[N],\, j\in[d],\, a_i\in\mathcal{A}_i} \big|\mathbb{E}_{h\sim f} g_{i,j,a_i}(h)\big| = 0.$$

Moreover, f is ϵ-decision calibrated if $\mathrm{DecCE}(f) \leq \epsilon$.

A decision-calibrated predictor ensures that receivers have no incentive to deviate from best responding to the prediction; that is, receivers cannot achieve higher utility by swapping their chosen action with another action. Similar guarantees have been established by Noarov et al. [2023], Roth and Shi [2024], and we provide a variant tailored to our setting of randomized predictor in the distributional setting. We formally define swap regret as follows:

Definition 2.4 (Swap Regret)**.** *We say that a predictor f achieves ϵ-swap regret if, for any receiver $i \in [N]$ and any mapping $\phi : \mathcal{A}_i \to \mathcal{A}_i$,*

$$\mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}\Big[\sum_{a_i\in\mathcal{A}_i} v_i(\phi(a), y)\, b_i(h(x), a_i)\Big] \leq \mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}\Big[\sum_{a_i\in\mathcal{A}_i} v_i(a, y)\, b_i(h(x), a_i)\Big] + \epsilon.$$

Theorem 2.5 (No Swap Regret via Decision Calibration)**.** *If a predictor f is ϵ-decision calibrated, then it has at most $2L|A|\epsilon$-swap regret.*

We further show that decision calibration can guarantee other forms of regret, such as *type regret*, which ensures that receivers have no incentive to pretend to be another receiver, as well as combinations of swap and type regret. These results demonstrate that decision calibration is a strongly compatible with receivers' incentive. We provide detailed statements and proofs in Appendix C.

We consider randomized predictors in $\Delta(\mathcal{H})$ that has decision calibration error bounded by some target level γ and assume that such predictors are not vacuous.

Assumption 2.6 (Feasibility). *There is a randomized predictor $f \in \Delta(\mathcal{H})$ that* $\mathrm{DecCE}(f) \leq \gamma$.

Note that Assumption 2.6 is mild since as long as $\mathcal{H}$ contains all deterministic constant predictors (or a discretized cover thereof), it necessarily includes a decision-calibrated predictor, specifically the constant predictor $h(x) = \mathbb{E}[Y]$, which is fully calibrated.

Sender's objective The sender does *not* have direct access to the data distribution $\mathcal{D}$, but has examples drawn from $\mathcal{D}$. The sender's utility depends on the outcome y and the joint action of all receivers, given by $u : \mathcal{A} \times \mathcal{Y} \to \mathbb{R}$ where $\mathcal{A} := \mathcal{A}_1 \times \cdots \times \mathcal{A}_N$. Without loss of generality, we assume the sender's utility is bounded by 1, i.e. $u(a, y) \in [0, 1]$ for any $a \in \mathcal{A}$ and $y \in \mathcal{Y}$. Given any prediction $h(x)$, for any $\boldsymbol{a} = (a_1, \cdots, a_N) \in \mathcal{A}$, the probability that receivers play joint action $\boldsymbol{a}$ is $b(h(x), \boldsymbol{a}) := \prod_{i=1}^{N} b_i(h(x), a_i)$.

The sender's goal is to maximize their expected utility subject to a γ-decision calibration constraint. Formally, the sender's optimization problem is

$$\max_{f} \mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y)b(h(x), \boldsymbol{a})\right] \quad \text{s.t.} \quad \mathrm{DecCE}(f) \leq \gamma. \tag{2}$$

We denote the optimal objective value of Eq. (2) by $\mathrm{OPT}(\mathcal{H}, \mathcal{D}, \gamma)$.

3 A minimax approach for efficient persuasive prediction

In this section, we present an efficient algorithm `PerDecCal` (Algorithm 1) for persuasive prediction, which learns an approximately optimal solution to the constrained optimization problem Eq. (2) from data when $|\mathcal{H}|$ is finite. We argue that finite hypothesis classes are already powerful: in finite-state Bayesian persuasion, a simple finite class suffices, as randomization over it can achieve the same sender utility as the full space of continuous signaling schemes. A detailed discussion is provided in Appendix B.

We begin by stating the theoretical guarantee achieved by `PerDecCal`.

Theorem 3.1. *Suppose `PerDecCal` runs for $T = O(\log(Nmd)/\epsilon^4)$ rounds and is given a dataset D drawn i.i.d. from $\mathcal{D}$ of size $n \geq O(\frac{\log(|\mathcal{H}|Ndm/\delta)}{\epsilon^4})$. With probability at least $1 - \delta$, the output predictor $\hat{f}$ that satisfies*

1. $\mathrm{DecCE}(\hat{f}) \leq \gamma + \epsilon$.
2. *Suppose the receivers play strict best response to $\hat{f}$. Then the receivers obtain swap regret bounded by $2mL(\gamma + \epsilon)$. The sender achieves ϵ-optimal utility:*

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}[u(\boldsymbol{a}, y) \cdot b(h(x), \boldsymbol{a})] \geq \mathrm{OPT}(\mathcal{H}, \mathcal{D}, \gamma) - \epsilon.$$

Theorem 3.1 shows that, with enough sample size, our proposed algorithm `PerDecCal` learns a predictor $\hat{f}$ that achieves nearly optimal utility compared to the best in-class γ-decision-calibrated predictor, while ensuring that its decision calibration error exceeds γ by at most ϵ, and this ensures that the receivers have no regret best responding to the predictions. It establishes a bi-criteria optimization: instead of requiring the predictor to be exactly γ-decision calibrated, we allow slight violations, which increase the swap regret by at most $2mL\varepsilon$. Furthermore, we prove a lower bound showing that it is statistically infeasible to learn a near-optimal predictor within the class of exactly γ-decision calibrated predictors. We defer the details to Appendix D.

`PerDecCal` follows a minimax-based approach. Specifically, we introduce Lagrangian variables and reformulate the original problem as a minimax game. We then apply an oracle-efficient algorithm to compute an approximate equilibrium of this game, which yields a near-optimal solution to the original problem Eq. (2). We now present the details and analysis of `PerDecCal`, with full proofs provided in Appendix E.

Lagrangian and C-bounded minimax game As a standard technique in optimization theory, the constrained optimization problem can be equivalently written in its Lagrangian form, which can be

interpreted as a minimax game. Specifically, we introduce the Lagrangian as follows:

$$\min_{f\in\Delta\mathcal{H}}\max_{\lambda\in\Lambda}\mathcal{L}_{\mathcal{D}}(f,\lambda) = -\mathbb{E}_{h\sim f,\mathcal{D}}[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)b(h(x),\boldsymbol{a})] + \sum_{s\in\{+,-\}}\sum_{i=1}^{N}\sum_{j=1}^{d}\sum_{a_i\in A_i}\lambda_{s,i,j,a_i}(s\cdot\mathbb{E}_{h\sim f}g_{i,j,a_i}(h)-\gamma). \tag{3}$$

Here, the dual variables are restricted to the bounded set $\lambda\in\Lambda=\{\lambda'|\lambda'\in\mathbb{R}_+^{2Nmd},\|\lambda\|_1\le C\}$.

Best response vs. no regret dynamics Viewing the problem as a minimax game, we consider solving it using *Best Response vs. No Regret* (BRNR) dynamics, where the min player f plays a best response to the current λ, and the max player updates λ according to a no-regret algorithm. Freund and Schapire [1996] showed that when both players achieve low regret, the average of their plays converges to an approximate equilibrium. Since both $\Delta(\mathcal{H})$ and Λ are convex spaces, we can apply their result once we establish sublinear regret for both players.

For the min player f, this is straightforward because $\mathcal{L}(f,\lambda)$ is linear in the randomized predictor f. As a result, the best response to any fixed λ is achieved by a deterministic predictor, that is, $\arg\min_{f\in\Delta(\mathcal{H})}\mathcal{L}(f,\lambda)=\arg\min_{h\in\mathcal{H}}\mathcal{L}(h,\lambda)$. To design a no-regret algorithm for the max player λ, we need the Lagrangian variables to be bounded, which is the reason why we consider $\lambda\in\Lambda$ instead of $\lambda\in\mathbb{R}_+^{2Nmd}$.

We next show that an approximate equilibrium (f,λ) of the C-bounded minimax game yields an approximately optimal solution to the original constrained problem in Eq. (2). Specifically, we prove that f achieves approximately optimal utility with respect to $\mathrm{OPT}(\mathcal{H},\mathcal{D},\gamma)$, while ensuring that its decision calibration error satisfies the γ-constraint up to an approximation error introduced by solving the C-bounded minimax game.

Lemma 3.2. *For an ϵ-approximate equilibrium of the C-bounded minimax game (f,λ). For the original unbounded constraint optimization problem Eq. (2), we have that $\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}[u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})]\ge\mathrm{OPT}(\mathcal{H},\mathcal{D},\gamma)-2\epsilon$, and $\mathrm{DecCE}(f)\le\gamma+\frac{1+2\epsilon}{C}$.*

Therefore, we reduce the sender's constrained optimization problem Eq. (2) to solving the equilibrium of the above bounded minimax game.

Solving bounded minimax games We now move on to solve the C-bounded minimax game. Note that the domain λ can be viewed as a scaling of the probability simplex. A natural choice of algorithm for this domain is a variant of the Hedge algorithm [Freund and Schapire, 1997], which is originally designed for the simplex. For simplicity, we scale the Hedge algorithm by C while still referring to it as Hedge. For computing the best response of the minimization player, we assume access to an empirical risk minimization (ERM) oracle that finds the best deterministic predictor given the current λ. We formally define the ERM oracle as follows:

Definition 3.3 (ERM oracle)**.** *Let the loss function be*

$$\ell_\lambda(h,x,y) := -\sum_{\boldsymbol{a}\in\mathcal{A}}u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a}) + \sum_{s\in\{+,-\}}\sum_{i=1}^{N}\sum_{j=1}^{d}\sum_{a_i\in A_i}\lambda_{s,i,j,a_i}(s\cdot\mathbb{E}_{h\sim f}g_{i,j,a_i}(h)-\gamma),$$

given a dataset with data points $D=\{(x_1,y_1),...,(x_n,y_n)\}$, the ERM oracle finds the best predictor that minimizes the empirical average loss: $\mathrm{ERM}(D,\lambda)=\arg\min_{h\in\mathcal{H}}\frac{1}{n}\sum_{i=1}^{n}\ell_\lambda(h,x_i,y_i)$.

The ERM oracle is commonly assumed in the learning theory literature and can often be implemented in practice using standard optimization methods. For example, when $\mathcal{H}$ is a class of neural networks, the ERM oracle can be approximated by running heuristic methods such as stochastic gradient descent (SGD) with a smoothed surrogate of the indicator $b(h(x),\boldsymbol{a})$ and $b_i(h(x),\boldsymbol{a})$.

We are now ready to present our algorithm `PerDecCal` in Algorithm 1. `PerDecCal` is efficient the ERM oracle is called only once per iteration, and all other operations are computationally polynomial in N,m,d. Therefore, overall `PerDecCal` is oracle-efficient, requiring $O(\log(Nmd)/\epsilon^4)$ calls to the ERM oracle.

Algorithm 1 `PerDecCal` (Persuasive Decision Calibration)

Input: A set of samples D, ERM oracle $\mathrm{ERM}(D, \lambda)$, dual bound C and tolerance γ.

1: Initialize $\lambda_1 = \frac{C}{2Nmd}\mathbf{1}$.
2: **for** $t = 1, \cdots, T$ **do**
3: Learner best responds to λ_t:
4: Use the ERM oracle to compute $h_t = \mathrm{ERM}(D, \lambda_t)$.
5: Auditor runs Hedge to obtain λ_{t+1}:
6: $\lambda_{t+1} = \mathrm{Hedge}(c_{1:t})$ where $c_t(\lambda_{s,i,j,a_i}) = \lambda_{s,i,j,a_i} s(\mathbb{E}_D[(h_t(x)_j - y_j) \cdot b_i(h_t(x), a_i)] - \gamma)$.
7: **end for**

Output: $\hat{f} = \mathrm{Uniform}(h_1, \cdots, h_T)$.

`PerDecCal` operates on the empirical dataset D instead of the true distribution $\mathcal{D}$. Therefore, a finite-sample analysis is needed to show that an approximate equilibrium found for $\mathcal{L}_D(f, \lambda)$ also serves as an approximate equilibrium for $\mathcal{L}_{\mathcal{D}}(f, \lambda)$. We prove a uniform convergence result showing that the payoff $\mathcal{L}_{\mathcal{D}}(f, \lambda)$ can be uniformly approximated by $\mathcal{L}_D(f, \lambda)$.

Lemma 3.4. *We have*

$$|\mathcal{L}_{\mathcal{D}}(f, \lambda) - \mathcal{L}_D(f, \lambda)| \leq \sqrt{\frac{\ln \frac{|4\mathcal{H}|}{\delta}}{2n}} + C\sqrt{\frac{8 \ln \frac{4|\mathcal{H}|Ndm}{\delta}}{n}}$$

for all $f \in \Delta(\mathcal{H}), \lambda \in \Lambda$ *with probability* $1 - \delta$.

The ϵ^{-4} dependence in Theorem 3.1 comes from setting $C = \Theta(1/\epsilon)$ for $O(\epsilon)$ calibration slack in Lemma 3.2. The average Hedge regret is then $\widetilde{O}(1/(\epsilon\sqrt{T}))$ and the uniform convergence error is $\widetilde{O}(1/(\epsilon\sqrt{n}))$. Making both errors $O(\epsilon)$ requires $\widetilde{O}(\epsilon^{-4})$ rounds and samples.

Then we can show that, with high probability, an approximate equilibrium under $\mathcal{L}_D(f, \lambda)$ is also an approximate equilibrium under $\mathcal{L}_{\mathcal{D}}(f, \lambda)$, completing the analysis of Theorem 3.1.

4 Matching the bayesian benchmark

In this section, we show that the sender utility achieved by our algorithm matches that of a fully informed Bayesian sender who is restricted to send signals induced by the same class of decision-calibrated predictors. Specifically, we compare against a restricted Bayesian persuasion benchmark in which both the sender and the receiver have full knowledge of the distribution $\mathcal{D}$, and the sender is constrained to commit to signaling schemes induced by $\mathcal{H}$ (defined in Definition 4.4), rather than all possible signaling schemes. Since the classical Bayesian persuasion model inherently involves a single receiver, we focus on the comparison within the single-receiver setting.

Bayesian persuasion benchmark The distribution $\mathcal{D}$ over $\mathcal{X} \times \mathcal{Y}$ which induces a distribution $\mu_{\mathcal{D}}$ over means of the outcome y conditional on the feature x: for any $\theta \in \mathcal{Y}$

$$\mu_{\mathcal{D}}(\theta) = \Pr_{(x,y)\sim\mathcal{D}}[\mathbb{E}[y \mid x] = \theta].$$

Here we use the fact that $\mathcal{Y}$ is convex. The set of state is $\Theta \subset \mathcal{Y}$ with prior $\mu_{\mathcal{D}}$. The receiver's action set $\mathcal{A}^{\mathrm{BP}}$ equals to the action set $\mathcal{A}$ in the prediction setting with utility function $v^{\mathrm{BP}} : \mathcal{A}^{\mathrm{BP}} \times \Theta \to \mathbb{R}$. The sender has utility function $u^{\mathrm{BP}} : \mathcal{A}^{\mathrm{BP}} \times \Theta \to \mathbb{R}$. We have $v^{\mathrm{BP}}(a, \theta) = \mathbb{E}_{y\sim\theta}[v(a, y)]$ and $u^{\mathrm{BP}}(a, \theta) = \mathbb{E}_{y\sim\theta}[u(a, y)]$. Here we slightly abuse notation by writing $y \sim \theta$ to indicate that y is drawn from a distribution with mean θ. A signaling scheme $\pi : \Theta \to \Delta(S)$ that randomly maps states to a set S of signals. Once the receiver observes a signal $s \in S$, they will update their belief from the prior μ to a posterior $\mu_s \in \Delta(\Theta)$ and consequently obtain a posterior mean that is in $\mathcal{Y}$. In other words, any signaling scheme will result in a distribution of posterior means $Q \in \Delta(\mathcal{Y})$.

We now establish the connection between decision calibration and the Bayesian persuasion benchmark introduced above. To do so, we use the notion of calibration as a bridge. Therefore, we begin by introduction the notion of calibration and presenting its relationship to Bayesian persuasion.

Definition 4.1 (Calibration)**.** *A randomized predictor* $f \in \Delta(\mathcal{H})$ *is said to be perfectly calibrated if*

$$\forall v \in \mathcal{Y}, \mathrm{CE}(f) := \mathbb{E}_{h\sim f,(x,y)\sim\mathcal{D}}[(y - h(x))|h(x) = v] = 0.$$

The following lemma states that every signaling scheme corresponds to a perfectly calibrated predictor, in the sense that the distribution over posterior means induced by the signaling scheme coincides with the distribution over predictions induced by a randomized calibrated predictor, and vice versa.

Lemma 4.2. *Consider a randomized predictor $f \in \Delta(\mathcal{H}_{\mathrm{ALL}})$ where $\mathcal{H}_{\mathrm{ALL}} = \{h : \mathcal{X} \to \mathcal{Y}\}$ is the class of all possible deterministic predictors. There exists a distribution $Q_f \in \Delta(\mathcal{Y})$ such that for any $v \in \mathcal{Y}$, $Q_f(v) = \Pr_{h\sim f,(x,y)\sim\mathcal{D}}[h(x) = v]$. A distribution $Q \in \Delta(\mathcal{Y})$ corresponds to the distribution over posterior means induced by some signaling scheme if and only if it is the prediction distribution Q_f of a perfectly calibrated predictor f.*

Note that the receiver's utility function v_i in the prediction setting is linear in the outcome y and the corresponding utility function v^{BP} in the Bayesian persuasion setting is linear in the conditional mean θ. By Lemma 4.2, it follows immediately that for the receiver, best responding to the predictions of a perfectly calibrated predictor is equivalent to best responding to the posterior means induced by the corresponding signaling scheme. Therefore, the calibrated predictor and the corresponding signaling scheme lead to the same sender's utility.

Now we present the connection between decision calibration and calibration. The next lemma says that any decision-calibrated predictor can be converted to a calibrated predictor without decreasing the sender's expected utility. Zhao et al. [2021] made a similar observation, though their result applies only to deterministic predictors, whereas we extend the analysis to randomized predictors.

Lemma 4.3. *For any randomized predictor f is perfectly decision calibrated, we can construct a randomized predictor f' such that (i) f' is perfectly calibrated; (ii) the sender obtains the same expected utility under f and f'.*

We are now ready to define the set of signaling schemes induced by $\mathcal{H}$ that we consider in our Bayesian persuasion benchmark.

Definition 4.4 (Signaling scheme class induced by $\mathcal{H}$)**.** *Given any class of deterministic predictors $\mathcal{H}$,*

1. *We define $\mathcal{F}_{\mathrm{DCAL}}(\mathcal{H})$ as the class of randomized predictor over $\mathcal{H}$ that is perfectly decision calibrated.*
2. *For any $f \in \mathcal{F}_{\mathrm{DCAL}}(\mathcal{H})$, let f' be the perfectly calibrated predictor constructed by Lemma 4.3. Define $\mathcal{F}_{\mathrm{CAL}}(\mathcal{H})$ as the class of all such predictors f'.*
3. *For any f' in $\mathcal{F}_{\mathrm{CAL}}(\mathcal{H})$, let $\pi_{f'}$ be the corresponding signaling scheme by Lemma 4.2. Define $\Pi_{\mathcal{H}}$ be the class of all such signaling schemes $\pi_{f'}$. We say that $\Pi_{\mathcal{H}}$ is the class of signaling schemes induced by $\mathcal{H}$.*

Finally we are ready to present our main result in this section.

Theorem 4.5. *Given at least $O(\frac{(\ln(|\mathcal{H}|dm/\delta)}{\epsilon^4})$ samples, with probability $1-\delta$, `PerDecCal` can output a predictor $\hat{f}$ such that the expected sender's utility under $\hat{f}$ is no worse than $\mathrm{BayesOPT}(\mu_{\mathcal{D}}, \Pi_{\mathcal{H}}) - \epsilon$ where we denote the optimal sender utility under our Bayesian persuasion benchmark as $\mathrm{BayesOPT}(\mu_{\mathcal{D}}, \Pi_{\mathcal{H}})$.*

Theorem 4.5 says the sender utility achieved by our algorithm matches that of a fully informed Bayesian sender who is restricted to send signals induced by the same class of predictors $\mathcal{H}$. Note that Theorem 4.5 naturally extends to the multi-receiver setting. In this case, the predictor can still match the Bayesian benchmark, where each receiver updates their posterior based on their own recommended action as the signal. However, if the receivers instead take the recommended joint action across all receivers as the signal, the decision calibration constraints need be modified to ensure unbiasedness with respect to the joint indicator $b(h(x), \boldsymbol{a})$ instead of individual $b_i(h(x), a)$.

Decision calibration is not the only constraint our framework can handle. The constrained learning problem must remain compatible with the optimization and auditing procedures in Section 3. Every feasible predictor must also admit a signaling scheme that is Bayes plausible and persuasive while preserving the sender's expected utility.

We provide a finite-hypothesis numerical validation of Theorem 4.5 in Appendix I, where the exact Bayesian benchmark can be computed by linear programming and our algorithm matches the corresponding optimum.

5 Persuasive prediction under infinite hypothesis class

In this section, we turn to the more general case where $|\mathcal{H}|$ can potentially be infinite. In this setting, strict best responses pose a challenge due to their discontinuity: even when two predictors $h_1, h_2 \in \mathcal{H}$ are close—i.e., $\sup_{x\in\mathcal{X}} \|h_1(x) - h_2(x)\|_\infty$ is small—the sender's utility under strict best response can differ by a constant. We illustrate this in the following example.

Example 5.1 (Discontinuity from best response)**.** *Consider a distribution $\mathcal{D}$ over $\mathcal{X} \times \mathcal{Y}$ such that $\mathcal{Y} = [0, 1]$ and for any $x \in \mathcal{X}$, $\Pr[y \mid x] = 0.5$. There is a single receiver, i.e. $N = 1$, who has two actions $\mathcal{A} = \{a, a'\}$ to choose from. The receiver has a utility function $v : \mathcal{A} \times \mathcal{Y} \to [0, 1]$ such that $v(a, y) = y, v(a', y) = 1 - y$. In other words, the receiver's best response is a when $y \le 0.5$ and is a' when $y > 0.5$. The sender has a utility function $v : \mathcal{A} \times \mathcal{Y} \to [0, 1]$ that only depends on the receiver's action: $u(a, y) = 1, u(a', y) = 0$. Now consider the following two predictors: for any $x \in \mathcal{X}$, $h_1(x) = 0.5 - \epsilon$ and $h_2(x) = 0.5 + \epsilon$. It is not hard to verify that both h_1 and h_2 are ϵ-decision calibrated predictor under distribution $\mathcal{D}$. However, for any $\epsilon \in (0, 0.5)$, the sender's expected utility is 1 under h_1, but 0 under h_2.*

Because of the discontinuity of the best response, even if $\mathcal{H}$ has a bounded complexity measure—such as a covering number or Rademacher complexity—it can be difficult to estimate the sender's utility (which best depends on receivers' best responses) uniformly over all $h \in \mathcal{H}$ from data. To overcome this challenge, we consider a *smoothed* version of the best response decision rule, commonly known as the *quantal response* model in economics and decision theory. This model has been extensively studied in the literature [McFadden et al., 1976, McKelvey and Palfrey, 1995] as it captures more realistic receiver behavior in the presence of noise, uncertainty, or bounded rationality. Unlike strict best responses, it allows receivers to probabilistically favor better actions while still occasionally choosing suboptimal ones, providing a smoother and more practical behavior model.

Definition 5.2 (Quantal Response)**.** *For any $i \in [N]$, the i-th receiver with utility function v_i responds to a prediction $h(x)$ according to the following η-*quantal response*:*

$$\tilde{b}_i(h(x), a_i) = \frac{e^{\eta v_i(a_i, h(x))}}{\sum_{a_i' \in \mathcal{A}_i} e^{\eta v_i(a_i', h(x))}}.$$

$\tilde{b}_i(h(x), a_i)$ denotes the probability that receiver i selects action a_i given the prediction $h(x)$. Here, $\eta > 0$ is the inverse temperature parameter, where as $\eta \to +\infty$, the receiver's behavior approaches the strict best response.

Analogously, we define a smoothed version of decision calibration when receivers follow quantal response model.

Definition 5.3 (Smoothed Decision Calibration)**.** *For a randomized predictor $f \in \Delta(\mathcal{H})$, define*

$$\tilde{g}_{i,j,a_i}(f) := \mathbb{E}_{h\sim f,\,(x,y)\sim\mathcal{D}}\left[(y_j - h(x)_j)\,\tilde{b}_i(h(x), a_i)\right].$$

We say that f is perfectly smoothed decision calibrated *if*

$$\mathrm{SmDecCE}(f) := \max_{i\in[N],\, j\in[d],\, a_i\in\mathcal{A}_i} \left|\tilde{g}_{i,j,a_i}(f)\right| = 0.$$

Moreover, f is ϵ-smoothed decision calibrated if $\mathrm{SmDecCE}(f) \le \epsilon$.

It can be shown that receivers have no regret when following the quantal response to a smoothed calibrated predictor; we refer the reader to Appendix C for a detailed discussion. Similar to Section 3, we design an oracle-effcient algorithm `SmPerDecCal` (Algorithm 2) for persuasive prediction given quantal response and $|\mathcal{H}|$ can be infinite.

Theorem 5.4. *Suppose `SmPerDecCal` runs for $T = O(\log(Nmd)/\epsilon^4)$ rounds and is given a dataset D drawn i.i.d of size $n \ge O(\ln \frac{\mathcal{N}(\mathcal{H}, d_\infty, \frac{\epsilon^2}{\eta L})Ndm}{\delta}/\epsilon^4)$. With probability at least $1 - \delta$, it outputs $\hat{f}$ that satisfies*

1. $\mathrm{SmDecCE}(\hat{f}) \le \gamma + \epsilon$.

2. *Suppose the receivers play η-quantal response to $\hat{f}$. Then the receivers obtain swap regret bounded by $2mL(\gamma+\epsilon)+\frac{\ln m+1}{\eta}$. The sender achieves ϵ-optimal utility:*

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\Big[u(\boldsymbol{a},y)\cdot\tilde{b}(h(x),\boldsymbol{a})\Big]\geq \mathrm{OPT}(\mathcal{H},\mathcal{D},\gamma)-\epsilon.$$

Theorem 5.4 shows that, with enough sample size, `SmPerDecCal` learns a predictor $\hat{f}$ that achieves nearly optimal utility compared to the best in-class γ-smoothed-decision-calibrated predictor, while ensuring that its smoothed decision calibration error exceeds γ by at most ϵ.

We empirically validate `SmPerDecCal` on the public FICO HELOC lending dataset, considering both a single-receiver setting and a multi-receiver variant with heterogeneous receiver preferences. In both settings, `SmPerDecCal` improves sender utility relative to cross-entropy ERM while maintaining comparable decision-calibration error. The full experimental setup and results are reported in Appendix H in the appendix.

Acknowledgments

We thank Dung Daniel Ngo for helpful discussions during the early stages of this project. JT, JZ, and ZSW are supported in part by an STTR grant. FF is supported in part by NSF grant IIS-2046640 (CAREER).

References

A. Agarwal, A. Beygelzimer, M. Dudík, J. Langford, and H. Wallach. A reductions approach to fair classification. In *International conference on machine learning*, pages 60–69. PMLR, 2018.

A. Agarwal, M. Dudík, and Z. S. Wu. Fair regression: Quantitative definitions and reduction-based algorithms. In *International conference on machine learning*, pages 120–129. PMLR, 2019.

G. Artemov, T. Kunimoto, and R. Serrano. Robust virtual implementation: Toward a reinterpretation of the wilson doctrine. *Journal of Economic Theory*, 148(2):424–447, 2013.

M.-F. Balcan, A. Blum, J. D. Hartline, and Y. Mansour. Reducing mechanism design to algorithm design via machine learning. *Journal of Computer and System Sciences*, 74(8):1245–1270, 2008.

M.-F. Balcan, A. Blum, N. Haghtalab, and A. D. Procaccia. Commitment without regrets: Online learning in stackelberg security games. In *Proceedings of the sixteenth ACM conference on economics and computation*, pages 61–78, 2015.

D. Bergemann and S. Morris. Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95, 2019. doi: 10.1257/jel.20181489.

M. K. Camara, J. D. Hartline, and A. Johnsen. Mechanisms for a no-regret agent: Beyond the common prior. In *2020 ieee 61st annual symposium on foundations of computer science (focs)*, pages 259–270. IEEE, 2020.

R. Cole and T. Roughgarden. The sample complexity of revenue maximization. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 243–252, 2014.

N. Collina, A. Roth, and H. Shao. Efficient prior-free mechanisms for no-regret agents. In *Proceedings of the 25th ACM Conference on Economics and Computation*, pages 511–541, 2024.

R. Cummings, N. R. Devanur, Z. Huang, and X. Wang. Algorithmic price discrimination. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2432–2451. SIAM, 2020.

C. Daskalakis and V. Syrgkanis. Learning in auctions: Regret is hard, envy is easy. In *2016 ieee 57th annual symposium on foundations of computer science (focs)*, pages 219–228. IEEE, 2016.

M. Dudík, N. Haghtalab, H. Luo, R. E. Schapire, V. Syrgkanis, and J. W. Vaughan. Oracle-efficient online learning and auction design. *Journal of the ACM (JACM)*, 67(5):1–57, 2020.

P. Dworczak and A. Pavan. Preparing for the worst but hoping for the best: Robust (bayesian) persuasion. *Econometrica*, 90(5):2017–2051, 2022.

C. Dwork, C. Hays, N. Immorlica, J. C. Perdomo, and P. Tankala. From fairness to infinity: Outcome-indistinguishable (omni) prediction in evolving graphs. *arXiv preprint arXiv:2411.17582*, 2024.

Y. Feng and W. Tang. Persuasive calibration. *arXiv preprint arXiv:2504.03211*, 2025.

FICO. Explainable machine learning challenge. https://community.fico.com/s/explainable-machine-learning-challenge, 2018.

M. Fishelson, R. Kleinberg, P. Okoroafor, R. P. Leme, J. Schneider, and Y. Teng. Full swap regret and discretized calibration. *arXiv preprint arXiv:2502.09332*, 2025.

D. P. Foster and R. Vohra. Regret in the on-line decision problem. *Games and Economic Behavior*, 29(1-2):7–35, 1999.

Y. Freund and R. E. Schapire. Game theory, on-line prediction and boosting. In *Proceedings of the ninth annual conference on Computational learning theory*, pages 325–332, 1996.

Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139, 1997.

H. Fu and T. Lin. Learning utilities and equilibria in non-truthful auctions. *Advances in Neural Information Processing Systems*, 33:14231–14242, 2020.

S. Garg, C. Jung, O. Reingold, and A. Roth. Oracle efficient online multicalibration and omniprediction. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2725–2792. SIAM, 2024.

I. Globus-Harris, V. Gupta, C. Jung, M. Kearns, J. Morgenstern, and A. Roth. Multicalibrated regression for downstream fairness. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, pages 259–286, 2023.

P. Gopalan, A. T. Kalai, O. Reingold, V. Sharan, and U. Wieder. Omnipredictors. *arXiv preprint arXiv:2109.05389*, 2021.

P. Gopalan, L. Hu, M. P. Kim, O. Reingold, and U. Wieder. Loss minimization through the lens of outcome indistinguishability. *arXiv preprint arXiv:2210.08649*, 2022.

P. Gopalan, P. Okoroafor, P. Raghavendra, A. Sherry, and M. Singhal. Omnipredictors for regression and the approximate rank of convex functions. In *The Thirty Seventh Annual Conference on Learning Theory*, pages 2027–2070. PMLR, 2024.

L. Hu and Y. Wu. Calibration error for decision making. *arXiv preprint arXiv:2404.13503*, 2024.

N. Immorlica, J. Mao, A. Slivkins, and Z. S. Wu. Incentivizing exploration with selective data disclosure. *arXiv preprint arXiv:1811.06026*, 2018.

A. Jain and V. Perchet. Calibrated forecasting and persuasion. *arXiv preprint arXiv:2406.15680*, 2024.

E. Kamenica and M. Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

M. Kearns, S. Neel, A. Roth, and Z. S. Wu. Preventing fairness gerrymandering: Auditing and learning for subgroup fairness. In *International conference on machine learning*, pages 2564–2572. PMLR, 2018.

B. Kleinberg, R. P. Leme, J. Schneider, and Y. Teng. U-calibration: Forecasting for an unknown agent. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 5143–5145. PMLR, 2023.

S. Kosterina. Persuasion with unknown beliefs. *Theoretical Economics*, 17(3):1075–1107, 2022.

T. Lin and C. Li. Information design with unknown prior. *arXiv preprint arXiv:2410.05533*, 2024.

J. Lu, A. Roth, and M. Shi. Sample efficient omniprediction and downstream swap regret for non-linear losses. *arXiv preprint arXiv:2502.12564*, 2025.

H. Luo, S. Senapati, and V. Sharan. Simultaneous swap regret minimization via kl-calibration. *arXiv preprint arXiv:2502.16387*, 2025.

D. McFadden et al. Quantal choice analysis: A survey. *Annals of economic and social measurement*, 5(4):363–390, 1976.

R. D. McKelvey and T. R. Palfrey. Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1):6–38, 1995.

J. H. Morgenstern and T. Roughgarden. On the pseudo-dimension of nearly optimal auctions. *Advances in Neural Information Processing Systems*, 28, 2015.

G. Noarov, R. Ramalingam, A. Roth, and S. Xie. High-dimensional prediction for sequential decision making. *arXiv preprint arXiv:2310.17651*, 2023.

P. Okoroafor, R. Kleinberg, and M. P. Kim. Near-optimal algorithms for omniprediction. *arXiv preprint arXiv:2501.17205*, 2025.

M. Ollár and A. Penta. Full implementation and belief restrictions. *American Economic Review*, 107 (8):2243–2277, 2017.

A. Parakhonyak and A. Sobolev. Persuasion without priors: A. parakhonyak, a. sobolev. *Economic Theory*, pages 1–35, 2025.

A. Roth and M. Shi. Forecasting for swap regret for all downstream agents. In *Proceedings of the 25th ACM Conference on Economics and Computation*, pages 466–488, 2024.

V. Syrgkanis. A sample complexity measure with applications to learning optimal auctions. *Advances in Neural Information Processing Systems*, 30, 2017.

J. Tang, J. Wu, Z. S. Wu, and J. Zhang. Dimension-free decision calibration for nonlinear loss functions. *arXiv preprint arXiv:2504.15615*, 2025.

S. Zhao, M. Kim, R. Sahoo, T. Ma, and S. Ermon. Calibrating predictions to decisions: A novel approach to multi-class calibration. *Advances in Neural Information Processing Systems*, 34: 22313–22324, 2021.

A Additional related work

Our work is conceptually related to recent work on prior-free mechanisms. Lin and Li [2024] studied Bayesian persuasion without knowing the prior, and showed that under certain regularity conditions it is possible to learn an approximately optimal signaling scheme by first estimating the prior from data and then solving the persuasion problem with the estimated distribution. Their approach is infeasible in our setting since we do not assume $\mathcal{D}$ to be learnable with a finite sample. Camara et al. [2020], Collina et al. [2024] study repeated Principal–Agent problems between long-lived principals and agents in adversarial, prior-free settings. To address the challenges arising from repeated interaction and online learning, these works introduce behavioral assumptions and machinery governing the agent's responses over time. Our work is complementary: we consider a statistically simpler batch persuasive-prediction setting with (smoothed) best-response receivers, which lets us optimize directly over covariate-dependent predictors subject to decision calibration. This yields in-class optimality result for richer prediction classes and vector-valued outcomes.

More broadly, our work shares the common goal of replacing prior knowledge with data, aligning with many works in mechanism design, such as auction design [Balcan et al., 2008, Cole and Roughgarden, 2014, Morgenstern and Roughgarden, 2015, Daskalakis and Syrgkanis, 2016, Syrgkanis, 2017, Dudík et al., 2020, Fu and Lin, 2020], Stackelberg game [Balcan et al., 2015], algorithm discrimination [Cummings et al., 2020] and recommendation system [Immorlica et al., 2018]. Relatedly, a growing body of work in economics aims to relax the assumption of perfect prior knowledge rather than replace it entirely, such as relaxing the prior to some form of approximate agreement on the distribution [Artemov et al., 2013, Ollár and Penta, 2017] and robustness to prior distribution [Dworczak and Pavan, 2022, Kosterina, 2022]. In contrast to these works, we adopt a data-driven approach to address the challenge of an unknown prior distribution. Parakhonyak and Sobolev [2025] also studies persuasion without a prior. Unlike our setting, which assumes sample access to the data distribution, they design a signaling scheme that minimizes the worst-case gap relative to the Bayesian benchmark.

Our work is related to a growing line of work on calibration in decision-making settings. The seminal work of Foster and Vohra [1999] showed that a decision maker who best responds to calibrated forecasts obtain diminishing internal regret. More recent studies extend this result by proposing refined notions of calibration that are more efficient to achieve and offer fine-grained regret guarantees [Zhao et al., 2021, Kleinberg et al., 2023, Hu and Wu, 2024, Roth and Shi, 2024, Fishelson et al., 2025, Luo et al., 2025, Tang et al., 2025]. Building on (multi)calibration, Gopalan et al. [2021] introduced the notion of omniprediction, which aims to construct a single predictor that guarantees no worse loss than a family of predetermined benchmarks for all the downstream receivers in a class, followed by Gopalan et al. [2022, 2024], Garg et al. [2024], Dwork et al. [2024], Okoroafor et al. [2025], Lu et al. [2025]. In contrast to these works, we not only aim to achieve a specific notion of calibration (decision calibration, in our case), but also seek to approximately maximize the sender's utility among all such calibrated predictors.

From a technical perspective, the problem of solving constrained optimization through No Regret versus Best Response dynamics has been studied in the algorithmic fairness literature [Agarwal et al., 2019, 2018, Kearns et al., 2018, Globus-Harris et al., 2023].

B Global optimality for finite-size $\mathcal{X}$

In this section, we consider the case that $|\mathcal{X}| < \infty$, $\mathcal{Y} = \{0, 1\}$ and there is one receiver, i.e. $N = 1$. We show that it is sufficient to consider predictions in an instance-dependent discretization set. Fix any receiver's utility v, we slightly abuse notation, let $\mathcal{A} = \{a_1, \cdots, a_m\}$ denote the receiver's action set. Let $J_i = \{p \in [0, 1] : a_i = \arg\max \mathbb{E}_{y\sim Ber(p)}[v(a, y)]\}$ for any $i \in [m]$. If there are ties, the receiver breaks ties in favor of the sender. Then we know that J_i is an interval on [0,1] for any $i \in [m]$ and $\{J_i\}_{i=1}^m$ is a partition of $[0, 1]$. Let $\mathcal{Z}$ be the set of thresholds of adjacent best response intervals. We know $|\mathcal{Z}| \leq m - 1$. Let $\Theta = \{\Pr[y = 1 \mid x] : \forall x \in \mathcal{X}\}$. We know $|\Theta| \leq |\mathcal{X}|$.

Definition B.1 (Instance-dependent discretization)**.** *For any positive $\epsilon < \min_{i\in m} \text{len}(J_i)$, define discretized set S_ϵ of space [0,1] as*

$$S_\epsilon \triangleq (\{0, \epsilon, 2\epsilon, \cdots\} \cap [0, 1]) \cup \mathcal{Z} \cup \Theta.$$

Definition B.2 (Discretized predictor set)**.** *Let $\mathcal{H}_\epsilon = \{h : \mathcal{X} \to S_\epsilon\}$ be the set of all possible predictors whose predictions are always in the instance-dependent discretization set S_ϵ.*

Note that $\mathcal{H}_\epsilon$ is finite with size $|S_\epsilon|^{\mathcal{X}}$. The following theorem shows that any randomized predictor in $\Delta(\mathcal{H}_{\text{ALL}})$ can be converted to a randomized predictor in $\Delta(\mathcal{H}_\epsilon)$ without changing the sender's utility and increase the decision calibration error up to ϵ.

Theorem B.3. *For any $\epsilon > 0$, for any randomized predictor $f \in \Delta(\mathcal{H}_{\text{ALL}})$ that is γ-decision calibrated, we can construct a randomized predictor $f' \in \Delta(\mathcal{H}_\epsilon)$ such that (1) the sender obtains the same expected utility under f and f' (2) f' is $(\gamma+\epsilon)$-decision calibrated.*

Proof. Define the range of f as $\text{range}(f) = \bigcup_{h\in\text{supp}\, f} \text{range}(h)$. Consider the following rounding function $r : \text{range}(f) \to S_\epsilon$:

$$r(v) = p \quad v \in J_i,\ p = \arg\inf_{p'\in(S_\epsilon\cap J_i)} |p' - v|.$$

For any $h \sim f \in \Delta(\mathcal{H}_{\text{ALL}})$, we construct $h' \in \mathcal{H}_\epsilon$ as $h' : x \mapsto r(h(x))$. Let f' be the distribution over such h'. By the definition of S_ϵ, we have that (1) $|v - r(v)| < \epsilon$ (2) the receiver has the same best response under v and $r(v)$. Therefore, the sender obtains the same expected utility under f and f'. And we have

$$\begin{aligned}
\left|\mathbb{E}_{h'\sim f'}\mathbb{E}_{(x,y)\sim\mathcal{D}}[(y-h(x))\cdot b(h'(x),a)]\right| &= \left|\mathbb{E}_{h'\sim f'}\mathbb{E}_{(x,y)\sim\mathcal{D}}[(y-h(x))\cdot b(h(x),a)]\right| \\
&\le \left|\mathbb{E}_{h'\sim f'}\mathbb{E}_{(x,y)\sim\mathcal{D}}[(y-h(x))\cdot b(h(x),a)]\right| \\
&\quad + \left|\mathbb{E}_{h'\sim f'}\mathbb{E}_{(x,y)\sim\mathcal{D}}[(h'(x)-h(x))\cdot b(h(x),a)]\right| \\
&= \left|\mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}[(y-h(x))\cdot b(h(x),a)]\right| \\
&\quad + \left|\mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}[(h'(x)-h(x))\cdot b(h(x),a)]\right| \\
&\le \gamma + \epsilon.
\end{aligned}$$

□

This theorem implies that, at least in the special case considered in this section, to find the optimal decision-calibrated predictor randomized over all deterministic predictors, it suffices to consider a finite subset of deterministic predictors $\mathcal{H}_\epsilon$.

C No regret guarantees of decision calibration

C.1 No regret guarantees of strict decision calibration

We first prove that approximate decision calibrated predictor gives the downstream agent no swap regret best responding to it. Noarov et al. [2023] proved it for deterministic predictor in the online calibration setting. We provide our proof for randomized predictor in the batch settting here for completeness.

Theorem 2.5 (No Swap Regret via Decision Calibration)**.** *If a predictor f is ϵ-decision calibrated, then it has at most $2L|A|\epsilon$-swap regret.*

Proof. We prove the result for any receiver $i \in [N]$.

$$\begin{aligned}
&\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a),y)\cdot b_i(h(x),a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,y)\cdot b_i(h(x),a)\right] \\
=&\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a),y)\cdot b_i(h(x),a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a),h(x))\cdot b_i(h(x),a)\right] \\
&+ \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a),h(x))\cdot b_i(h(x),a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot b_i(h(x),a)\right] \\
&+ \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot b_i(h(x),a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,y)\cdot b_i(h(x),a)\right]
\end{aligned}$$

When f is ϵ-decision-calibrated, we know that

$$\begin{aligned}
&\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a),y)\cdot b_i(h(x),a)\right]-\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a),h(x))\cdot b_i(h(x),a)\right]\\
=&\sum_a \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a),y-h(x))\cdot b_i(h(x),a)\right]\\
\leq&\sum_a L\epsilon = L|A|\epsilon.
\end{aligned}$$

Similarly we can prove that

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot b_i(h(x),a)\right]-\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,y)\cdot b_i(h(x),a)\right]\leq L|A|\epsilon.$$

Since $b_i(h(x),a)$ plays the best response given the prediction $h(x)$, we know that

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a),h(x))\cdot b_i(h(x),a)\right]-\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot b_i(h(x),a)\right]\leq 0.$$

Putting them together, we have

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a),y)\cdot b_i(h(x),a)\right]-\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,y)\cdot b_i(h(x),a)\right]\leq 2L|A|\epsilon.$$

□

Definition C.1 (Type Regret)**.** *We say that a predictor f achieves ϵ-type regret if, for any receiver $i,i'\in[N]$,$\phi:A\to A$,*

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,y)\cdot b_{i'}(h(x),a)\right]\leq\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,y)\cdot b_i(h(x),a)\right]+\epsilon.$$

Now we introduce a different notion of regret, named *type regret*. Type regret is first introduced by Zhao et al. [2021]. Intuitively, it says that once the predictor gets decision calibrated with respect to a class of utility functions of receivers, the receivers will have no regret best responding according to another receiver's utility function instead their own. Zhao et al. [2021] proved that decision calibrated predictor achieves no type regret for the receivers. Here we state and prove the result for randomized predictors.

Theorem C.2 (No Type Regret via Decision Calibration)**.** *If a predictor f is ϵ-decision calibrated, then it satisfies $2L|A|\epsilon$-type regret.*

Proof. We prove the result for any receiver $i,i'\in[N]$.

$$\begin{aligned}
&\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,y)\cdot b_{i'}(h(x),a)\right]-\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,y)\cdot b_i(h(x),a)\right]\\
=&\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,y)\cdot b_{i'}(h(x),a)\right]-\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot b_{i'}(h(x),a)\right]\\
&+\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot b_{i'}(h(x),a)\right]-\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot b_i(h(x),a)\right]\\
&+\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot b_i(h(x),a)\right]-\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,y)\cdot b_i(h(x),a)\right]
\end{aligned}$$

When f is ϵ-decision-calibrated, we know that

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,y)\cdot b_{i'}(h(x),a)\right]-\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot b_{i'}(h(x),a)\right]$$

$$= \sum_a \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, y - h(x)) \cdot b_{i'}(h(x), a)\right]$$
$$\leq \sum_a L\epsilon = L|A|\epsilon.$$

Similarly we can prove that

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, h(x)) \cdot b_{i'}(h(x), a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, h(x)) \cdot b_i(h(x), a)\right] \leq L|A|\epsilon.$$

Since $b_i(h(x), a)$ plays the best response given the prediction $h(x)$, we know that

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, h(x)) \cdot b_{i'}(h(x), a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, h(x)) \cdot b_i(h(x), a)\right] \leq 0.$$

Putting them together, we have

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, y) \cdot b_{i'}(h(x), a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, y) \cdot b_i(h(x), a)\right] \leq 2L|A|\epsilon.$$

□

We also introduce a new notion of regret which we call swap-type-regret, which intuitively capture the case where the receivers can first pretend that they were another receiver and then swap the corresponding best-response action. We formally define it as follows:

Definition C.3 (Swap-Type Regret)**.** *We say that a predictor f achieves ϵ-swap-type regret if, for any receiver $i, i' \in [N]$, mapping function $\phi : A \to A$,*

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a), y) \cdot b_{i'}(h(x), a)\right] \leq \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, y) \cdot b_i(h(x), a)\right] + \epsilon.$$

Theorem C.4 (No Swap-Type Regret via Decision Calibration)**.** *If a predictor f is ϵ-decision calibrated, then it satisfies $2L|A|\epsilon$-swap-type regret.*

Proof. The proof is similar to the proofs of Theorem 2.5 and Theorem C.2. We can similarly prove that

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a), y) \cdot b_{i'}(h(x), a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a), h(x)) \cdot b_{i'}(h(x), a)\right] \leq L|A|\epsilon,$$

and

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, h(x)) \cdot b_i(h(x), a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, y) \cdot b_i(h(x), a)\right] \leq L|A|\epsilon.$$

From the fact that $b_i(h(x), a)$ selects the best response action, we also have

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a), h(x)) \cdot b_{i'}(h(x), a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, h(x)) \cdot b_i(h(x), a)\right] \leq 0.$$

Putting them together completes the proof. □

C.2 No regret guarantees of smoothed decision calibration

We provide analogous result for the behavior model where the receivers follow quantal response. We first define the three notions of regret for quantal response.

Definition C.5 (Swap Regret under Quantal Response)**.** *We say that a predictor f achieves ϵ-swap regret for receivers that follow quantal response rule if, for any receiver $i \in [N]$, mapping function $\phi : A \to A$,*

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a), y) \cdot \tilde{b}_i(h(x), a)\right] \leq \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, y) \cdot \tilde{b}_i(h(x), a)\right] + \epsilon.$$

Definition C.6 (Type Regret under Quantal Response)**.** *We say that a predictor f achieves ϵ-type regret for receivers that follow quantal response rule if, for any receiver $i, i' \in [N]$,$\phi : A \to A$,*

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, y) \cdot \tilde{b}_{i'}(h(x), a)\right] \leq \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, y) \cdot \tilde{b}_i(h(x), a)\right] + \epsilon.$$

Definition C.7 (Swap-Type under Quantal Response)**.** *We say that a predictor f achieves ϵ-swap-type regret for receivers that follow quantal response rule if, for any receiver $i, i' \in [N]$, mapping function $\phi : A \to A$,*

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a), y) \cdot \tilde{b}_{i'}(h(x), a)\right] \leq \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, y) \cdot \tilde{b}_i(h(x), a)\right] + \epsilon.$$

We now present the analogous no regret guarantee for quantal response receivers. Swap regret for quantal response receivers are discussed in Roth and Shi [2024] in the online setting for deterministic predictors. Here, we provide the result for randomized predictors in the batch setting. Type regret for quantal response receivers are discussed in Tang et al. [2025] for determinisitc predictors. Here, we state and prove the result for randomized predictors.

Theorem C.8. *If a predictor f is ϵ-smoothed-decision calibrated, then it satisfies $2L|A|\epsilon + \frac{\ln|A|+1}{\eta}$-swap/type/swap-type regret.*

To prove Theorem C.8, we first present a lemma proved by Roth and Shi [2024], which shows that the utility that a receiver gets when they play quantal response with respect to the true outcome will be close that when they strictly best responds.

Lemma C.9 (Roth and Shi [2024])**.** *For any utility function v and $y \in \mathcal{Y}$, let $a^* = \arg\max_a v(a, y)$, we have*

$$\sum_a v(a, y)\tilde{b}(y, a) \geq v(a^*, y) - \frac{\ln|A| + 1}{\eta}.$$

Proof of Theorem C.8. We only provide proof for swap-type regret as it is the strongest. The guarantees for swap regret and type regret are simple corollaries by considering ϕ to be the identical mapping and $i = i'$.

For any $\phi : \mathcal{A} \times \mathcal{A}, i, i' \in [N]$, we have

$$\begin{aligned}
&\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a), y) \cdot \tilde{b}_{i'}(h(x), a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, y) \cdot \tilde{b}_i(h(x), a)\right] \\
\leq&\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a), y) \cdot \tilde{b}_{i'}(h(x), a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a), h(x)) \cdot \tilde{b}_{i'}(h(x), a)\right] \\
&+ \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a), h(x)) \cdot \tilde{b}_{i'}(h(x), a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, h(x)) \cdot \tilde{b}_i(h(x), a)\right] \\
&+ \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, h(x)) \cdot \tilde{b}_i(h(x), a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a, y) \cdot \tilde{b}_i(h(x), a)\right]
\end{aligned}$$

From the definition of decision calibration, we know that

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a), y) \cdot \tilde{b}_{i'}(h(x), a)\right] - \mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a), h(x)) \cdot \tilde{b}_{i'}(h(x), a)\right] \leq L|A|\epsilon$$

and

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot\tilde{b}_i(h(x),a)\right]-\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,y)\cdot\tilde{b}_i(h(x),a)\right]\le L|A|\epsilon.$$

By Lemma C.9, we know that

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot\tilde{b}_i(h(x),a)\right]\ge\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot b_i(h(x),a)\right]-\frac{\ln|A|+1}{\eta}.$$

Therefore, we have

$$\begin{aligned}&\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a),h(x))\cdot\tilde{b}_{i'}(,h(x),a)\right]-\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot\tilde{b}_i(h(x),a)\right]\\ \le&\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a),h(x))\cdot b_{i'}(,h(x),a)\right]\\ &-\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot b_i(,h(x),a)\right]+\frac{\ln|A|+1}{\eta}.\end{aligned}$$

Since given $h(x)$, $b_i(h(x),a)$ is the optimal decision rule for v_i, we know that

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a),h(x))\cdot b_{i'}(h(x),a)\right]-\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,h(x))\cdot b_i(h(x),a)\right]\le 0.$$

Putting them together, we have that

$$\begin{aligned}&\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(\phi(a),y)\cdot b_{i'}(h(x),a)\right]-\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\left[\sum_a v_i(a,y)\cdot b_i(h(x),a)\right]\\ \le& 2L|A|\epsilon+\frac{\ln|A|+1}{\eta}.\end{aligned}$$

□

D Statistical hardness of learning an optimal predictor with the exact constraints

In this section, we show that it is statistically hard to learn the optimal predictor in the class of γ-decision calibrated predictor. Specially, we focus on $\gamma=0$. Then an algorithm is called (ε,δ)-bicriteria optimal if, there exist a function $n_0:(0,1)^2\to\mathbb{N}$ such that given $n\ge n_0(\epsilon,\delta)$ i.i.d. samples from an unknown distribution $\mathcal{D}$, it outputs a (possibly randomized) predictor f such that, with probability at least $1-\delta$,

$$\text{(i) } \mathrm{DecCE}(f)\le\varepsilon \qquad \text{and} \qquad \text{(ii) } \mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}\Big[\sum_{a\in\mathcal{A}}u(a,y)\,b(h(x),a)\Big]\ \ge\ \mathrm{OPT}(\mathcal{H},\mathcal{D},\gamma)-\varepsilon.$$

Instance. Consider one receiver (N=1), outcomes $y\in\{0,1\}$, a feature space $\mathcal{X}=\{0\}$, and a hypothesis class

$$\mathcal{H}=\{h_1,h_2\},\qquad h_1(x)\equiv v_1,\ h_2(x)\equiv v_2,\quad 0<v_1<v_2<1.$$

Suppose the receiver's optimal action is a_1 if $y\ge v_1$ and a_2 otherwise. The sender receives utility 1 if the receiver chooses a_1, and 0 otherwise. Let $\mathcal{D}_-$ be the distribution with $x\equiv 0$ and $y\sim\mathrm{Bern}(v_1)$, and $\mathcal{D}_+$ the same with $y\sim\mathrm{Bern}(v_2)$. Define $\Delta:=v_2-v_1=\mathrm{TVD}(\mathcal{D}_-,\mathcal{D}_+)$ and assume $v_1,v_2\in[\tau,1-\tau]$ for some $\tau\in(0,\frac{1}{2})$. Choose ε so that $0<\varepsilon<\Delta/2$.

For any (possibly randomized) $f\in\Delta(\mathcal{H})$,

$$\mathrm{DecCE}(f)=\Big|\,\mathbb{E}_{h\sim f}\big[\mathbb{E}[y]-h(0)\big]\,\Big|. \tag{4}$$

Theorem D.1. *For the instance above, any (ε, δ)-bicriteria algorithm with $0 < \varepsilon < \Delta/2$ and $0 < \delta < \frac{1}{2}$ that succeeds using n samples must satisfy*

$$n \;\geq\; \frac{2\,\tau(1-\tau)\,(1-2\delta)^2}{\Delta^2}.$$

Equivalently, the sample complexity is $\Omega((1-2\delta)^2/\mathrm{TVD}(\mathcal{D}_-, \mathcal{D}_+)^2)$.

Proof. From equation 4, under $\mathcal{D}_-$ the feasibility condition $\mathrm{DecCE}(f) \leq \varepsilon$ implies $\mathbb{E}_{h\sim f}[h(0)] \in [v_1 - \varepsilon,\, v_1 + \varepsilon]$; under $\mathcal{D}_+$ it implies $\mathbb{E}_{h\sim f}[h(0)] \in [v_2 - \varepsilon,\, v_2 + \varepsilon]$. These intervals are disjoint because $2\varepsilon < \Delta$.

Under $\mathcal{D}_-$ (resp. $\mathcal{D}_+$), the exactly calibrated point mass δ_{h_1} (resp. δ_{h_2}) is optimal, i.e., $\mathrm{OPT}(\mathcal{H}, \mathcal{D}_\pm, 0) = 1$. Moreover, any $f \in \Delta(\mathcal{H})$ with $\mathrm{DecCE}(f) \leq \varepsilon$ is also ε-optimal for the objective.

Let an algorithm be (ε, δ)-bicriteria optimal. With probability at least $1 - \delta$ it outputs an f that is ε-feasible (and therefore ε-optimal). Consequently, the statistic $\mu := \mathbb{E}_{h\sim f}[h(0)]$ lies in disjoint intervals depending on whether the data come from $\mathcal{D}_-$ or $\mathcal{D}_+$. Thus declaring $\mathcal{D}_+$ iff $\mu > (v_1+v_2)/2$ identifies the generating distribution with error at most δ.

For any rule that distinguishes $\mathcal{D}_-^{\otimes n}$ from $\mathcal{D}_+^{\otimes n}$ with error at most δ one must have $\mathrm{TVD}(\mathcal{D}_-^{\otimes n},\, \mathcal{D}_+^{\otimes n}) \geq 1 - 2\delta$. Pinsker's inequality and the KL chain rule give

$$\mathrm{TVD}(\mathcal{D}_-^{\otimes n},\, \mathcal{D}_+^{\otimes n}) \;\leq\; \sqrt{\frac{n}{2}\, D_{\mathrm{KL}}(\mathcal{D}_- \,\|\, \mathcal{D}_+)}.$$

For Bernoulli parameters $p, q \in [\tau, 1-\tau]$, $D_{\mathrm{KL}}(\mathrm{Bern}(p) \,\|\, \mathrm{Bern}(q)) \leq (p-q)^2/(\tau(1-\tau))$. Here $|p - q| = \Delta$, so

$$\mathrm{TVD}(\mathcal{D}_-^{\otimes n},\, \mathcal{D}_+^{\otimes n}) \;\leq\; \sqrt{\frac{n\,\Delta^2}{2\,\tau(1-\tau)}}.$$

Combining with $\mathrm{TVD} \geq 1 - 2\delta$ yields $\sqrt{n\,\Delta^2/(2\,\tau(1-\tau))} \geq 1 - 2\delta$, i.e., $n \geq 2\,\tau(1-\tau)\,(1-2\delta)^2/\Delta^2$.

□

With zero slack ($\epsilon = 0$) the condition $\epsilon < \Delta/2$ holds for any $\Delta > 0$; since we can choose $\Delta = v_2 - v_1$ arbitrarily small while keeping $v_1, v_2 \in [\tau, 1-\tau]$, Theorem D.1 gives $n \geq 2\tau(1-\tau)(1-2\delta)^2/\Delta^2 \to \infty$ as $\Delta \to 0$, i.e., the sample complexity is unbounded.

E Missing Proofs in Section 3

We first state the lemma from Freund and Schapire [1996] which proves that when both players have low regret, their average play converge to an approximate equilibrium.

Lemma E.1 (Freund and Schapire [1996])**.** *Consider a two-player zero-sum game where the min player chooses strategies from $\mathcal{P}$ and the max player chooses strategies from $\mathcal{Q}$. Assume $\mathcal{P}$ and $\mathcal{Q}$ are convex, and the utility function is bilinear in the players' strategies. If the sequence of plays satisfies sublinear regret for both players, i.e.,*

$$\min_{p\in\mathcal{P}} \sum_{t=1}^{T} u(p_t, q_t) - u(p, q_t) \leq \gamma_{\mathcal{P}} T, \quad \textit{and} \quad \max_{q\in\mathcal{Q}} \sum_{t=1}^{T} u(p_t, q) - u(p_t, q_t) \leq \gamma_{\mathcal{Q}} T,$$

then letting $\bar{p} = \frac{1}{T}\sum_{t=1}^{T} p_t$ and $\bar{q} = \frac{1}{T}\sum_{t=1}^{T} q_t$, we have that $(\bar{p}, \bar{q})$ is a $(\gamma_{\mathcal{P}} + \gamma_{\mathcal{Q}})$-approximate minimax equilibrium of the game.

Lemma 3.2. *For an ϵ-approximate equilibrium of the C-bounded minimax game (f, λ). For the original unbounded constraint optimization problem Eq. (2), we have that $\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}[u(\boldsymbol{a}, y)\cdot b(h(x), \boldsymbol{a})] \geq \mathrm{OPT}(\mathcal{H}, \mathcal{D}, \gamma) - 2\epsilon$, and $\mathrm{DecCE}(f) \leq \gamma + \frac{1+2\epsilon}{C}$.*

Proof. We use f^* to denote the optimal feasible solution to Eq. (2), since the constraints are satisfied, we have that $\mathcal{L}(f^*, \hat{\lambda}) \leq \mathrm{OPT}(\mathcal{H}, \mathcal{D}, \gamma)$. We prove the theorem by considering two cases.

First, if $\hat{f}$ is a feasible solution to the problem Eq. (2), i.e. $\mathrm{DecCE}(\hat{f}) \leq \gamma \leq \gamma + \frac{1+2\epsilon}{C}$. Since $(\hat{f}, \bar{\lambda})$ is a ϵ-approximate equilibrium, we have that

$$\begin{aligned}
-\mathbb{E}_{h\sim f}\mathbb{E}_D[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y)\cdot b(h(x), \boldsymbol{a})] &= \max_{\lambda} \mathcal{L}(\hat{f}, \lambda) \\
&\leq \mathcal{L}(\hat{f}, \bar{\lambda}) + \epsilon \\
&\leq \min_{f\in\Delta(\mathcal{H})} \mathcal{L}(\hat{f}, \bar{\lambda}) + 2\epsilon \\
&\leq \mathcal{L}(f^*, \bar{\lambda}) + 2\epsilon \\
&\leq -\mathbb{E}_{h\sim f^*}\mathbb{E}_D[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y)\cdot b(h(x), \boldsymbol{a})] + 2\epsilon \\
&= -\mathrm{OPT}(\mathcal{H}, \mathcal{D}, \gamma).
\end{aligned}$$

Therefore, we have

$$\mathbb{E}_{h\sim f}\mathbb{E}_D[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y)\cdot b(h(x), \boldsymbol{a})] \geq \mathrm{OPT}(\mathcal{H}, \mathcal{D}, \gamma) - 2\epsilon.$$

Second, we consider the case where $\hat{f}$ is not a feasible solution to Eq. (2). Let $(\hat{s}, \hat{i}, \hat{j}, \hat{a}_i) = \arg\max_{s,i,j,a_i} s(\mathbb{E}_f\mathbb{E}_D[(h(x)_j - y_j)\cdot b_i(h(x), a_i)] - \gamma) > 0$, and let λ' be the vector such that the $(\hat{s}, \hat{i}, \hat{j}, \hat{a}_i)$-th coordinate $\lambda'_{\hat{s},\hat{i},\hat{j},\hat{a}_i} = C$, and all else coordinates are 0, then we have that given $\lambda' \in \arg\max_\lambda \mathcal{L}(\hat{f}, \lambda)$. Therefore, since $(\hat{f}, \bar{\lambda})$ is a ϵ-approximate equilibrium, we have

$$\begin{aligned}
&\mathcal{L}(\hat{f}, \bar{\lambda}) \geq \max_{\lambda} \mathcal{L}(\hat{f}, \lambda) - \epsilon \\
&\quad = -\mathbb{E}_{h\sim f}\mathbb{E}_D[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y)\cdot b(h(x), \boldsymbol{a})] + C\hat{s}\Big(\mathbb{E}_f\mathbb{E}_D[(h(x)_{\hat{j}} - y_{\hat{j}})\cdot b_{\hat{i}}(h(x), a_{\hat{i}})] - \gamma\Big) - \epsilon
\end{aligned}$$

Therefore,

$$\begin{aligned}
&-\mathbb{E}_{h\sim f}\mathbb{E}_D[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y)\cdot b(h(x), \boldsymbol{a})] + C\hat{s}\Big(\mathbb{E}_f\mathbb{E}_D[(h(x)_{\hat{j}} - y_{\hat{j}})\cdot b_{\hat{i}}(h(x), a_{\hat{i}})] - \gamma\Big) \\
&\leq \mathcal{L}(\hat{f}, \bar{\lambda}) + \epsilon \\
&\leq \mathcal{L}(f^*, \bar{\lambda}) + 2\epsilon \\
&\leq -\mathbb{E}_{h\sim \hat{f}}\mathbb{E}_D[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y)\cdot b(h(x), \boldsymbol{a})] + 2\epsilon
\end{aligned}$$

Since $\forall \boldsymbol{a}, y, u(\boldsymbol{a}, y) \in [0, 1]$, we have that

$$C\hat{s}\Big(\mathbb{E}_f\mathbb{E}_D[(h(x)_{\hat{j}} - y_{\hat{j}})\cdot b_{\hat{i}}(h(x), a_{\hat{i}})] - \gamma\Big) \leq 1 + 2\epsilon.$$

Thus,

$$\max_{s,i,j,a_i} s(\mathbb{E}_f\mathbb{E}_D[(h(x)_j - y_j)\cdot b_i(h(x), a_i)] - \gamma) \leq \frac{1+2\epsilon}{C},$$

and this implies that $\mathrm{DecCE}(\hat{f}) \leq \gamma + \frac{1+2\epsilon}{C}$. □

We need the following technical lemmas before proving Lemma 3.4. For any function $\psi : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}$ and any dataset $D = \{(x^{(i)}, y^{(i)})\}_{i=1}^n$, we denote the empirical expectation of ψ over D as

$$\hat{\mathbb{E}}_D[\psi(x, y)] \triangleq \frac{1}{n}\sum_{i=1}^n \psi(x^{(i)}, y^{(i)}).$$

Theorem E.2. *Fix a finite-size class of deterministic predictors $\mathcal{H}$. For any distribution $\mathcal{D}$, let $D \sim \mathcal{D}^n$ be a dataset consisting of n samples $(x^{(i)}, y^{(i)})$ sampled i.i.d. from $\mathcal{D}$. Then for any $\delta \in (0,1)$, with probability $1-\delta$, for every $f \in \Delta(\mathcal{H})$, we have*

$$\left|\mathbb{E}_{\mathcal{D}}\mathbb{E}_{h\sim f}\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})\right] - \hat{\mathbb{E}}_D\mathbb{E}_{h\sim f}\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})\right]\right| \le \sqrt{\frac{\ln\frac{2|\mathcal{H}|}{\delta}}{2n}}.$$

Proof. For any $(x,y) \sim \mathcal{D}$, observe that $\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a}) \le \sum b(h(x),\boldsymbol{a}) = 1$. By Hoeffding's inequality, we have for any $h \in \mathcal{H}$, for any $\delta' \in (0,1)$ with probability $1-\delta'$, we have

$$\left|\mathbb{E}_{\mathcal{D}}\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})\right] - \hat{\mathbb{E}}_D\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})\right]\right| \le \sqrt{\frac{\ln\frac{2}{\delta'}}{2n}}.$$

Then let $\delta' = \delta/|\mathcal{H}|$, by union bound we have with probability $1-\delta$, for any $h \in \mathcal{H}$,

$$\left|\mathbb{E}_{\mathcal{D}}\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})\right] - \hat{\mathbb{E}}_D\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})\right]\right| \le \sqrt{\frac{\ln\frac{2|\mathcal{H}|}{\delta}}{2n}}.$$

Finally we have with probability $1-\delta$, for any $f \in \Delta(\mathcal{H})$,

$$\begin{aligned}
&\left|\mathbb{E}_{\mathcal{D}}\mathbb{E}_{h\sim f}\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})\right] - \hat{\mathbb{E}}_D\mathbb{E}_{h\sim f}\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})\right]\right| \\
&= \left|\mathbb{E}_{h\sim f}\left[\mathbb{E}_{\mathcal{D}}\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})\right]\right] - \mathbb{E}_{h\sim f}\left[\hat{\mathbb{E}}_D\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})\right]\right]\right| \\
&= \left|\mathbb{E}_{h\sim f}\left[\mathbb{E}_{\mathcal{D}}\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})\right] - \hat{\mathbb{E}}_D\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})\right]\right]\right| \\
&\le \mathbb{E}_{h\sim f}\left[\left|\mathbb{E}_{\mathcal{D}}\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})\right] - \hat{\mathbb{E}}_D\left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})\right]\right|\right] \\
&\le \mathbb{E}_{h\sim f}\left[\sqrt{\frac{\ln\frac{2|\mathcal{H}|}{\delta}}{2n}}\right] = \sqrt{\frac{\ln\frac{2|\mathcal{H}|}{\delta}}{2n}}.
\end{aligned}$$

□

Theorem E.3. *Fix a finite-size class of deterministic predictors $\mathcal{H}$. For any distribution $\mathcal{D}$, let $D \sim \mathcal{D}^n$ be a dataset consisting of n samples $(x^{(i)}, y^{(i)})$ sampled i.i.d. from $\mathcal{D}$. Then for any $\delta \in (0,1)$, with probability $1-\delta$, for every $f \in \Delta(\mathcal{H}), i \in [N], j \in [d], a_i \in \mathcal{A}_i$ we have*

$$\begin{aligned}
&\left|\mathbb{E}_{\mathcal{D}}\mathbb{E}_{h\sim f}[(y_j - h(x)_j)\cdot b_i(h(x),a_i)] - \hat{\mathbb{E}}_D\mathbb{E}_{h\sim f}[(y_j - h(x)_j)\cdot b_i(h(x),a_i)]\right| \\
&\le \sqrt{\frac{8\ln\frac{2|\mathcal{H}|Ndm}{\delta}}{n}}.
\end{aligned}$$

Proof. For any $(x,y) \sim \mathcal{D}, i \in [N], j \in [d]$, observe that $(y_j - h(x)_j)\cdot b_i(h(x),a_i) \in [-2,2]$. By Hoeffding's inequality, we have for any $h \in \mathcal{H}$, for any $\delta' \in (0,1)$ with probability $1-\delta'$, we have

$$\left|\mathbb{E}_{\mathcal{D}}[(y_j - h(x)_j)\cdot b_i(h(x),a_i)] - \hat{\mathbb{E}}_D[(y_j - h(x)_j)\cdot b_i(h(x),a_i)]\right| \le \sqrt{\frac{8\ln\frac{2}{\delta'}}{n}}.$$

Then let $\delta' = \delta/(|\mathcal{H}|Ndm)$, by union bound we have with probability $1-\delta$, for any $h \in \mathcal{H}, i \in [N]$ and $j \in [d]$ and $a_i \in \mathcal{A}_i$

$$\left|\mathbb{E}_{\mathcal{D}}[(y_j - h(x)_j)\cdot b_i(h(x),a_i)] - \hat{\mathbb{E}}_D[(y_j - h(x)_j)\cdot b_i(h(x),a_i)]\right| \le \sqrt{\frac{8\ln\frac{2|\mathcal{H}|Ndm}{\delta}}{n}}.$$

Finally we have with probability $1-\delta$, for any $f \in \Delta(\mathcal{H})$,

$$\begin{aligned}
&\left|\mathbb{E}_{(x,y)\sim\mathcal{D}}\mathbb{E}_{h\sim f}[(y_j - h(x)_j)\cdot b_i(h(x),a_i)] - \hat{\mathbb{E}}_D\mathbb{E}_{h\sim f}[(y_j - h(x)_j)\cdot b_i(h(x),a_i)]\right| \\
&= \left|\mathbb{E}_{h\sim f}\left[\mathbb{E}_{(x,y)\sim\mathcal{D}}[(y_j - h(x)_j)\cdot b_i(h(x),a_i)]\right] - \mathbb{E}_{h\sim f}\left[\hat{\mathbb{E}}_D[(y_j - h(x)_j)\cdot b_i(h(x),a_i)]\right]\right| \\
&= \left|\mathbb{E}_{h\sim f}\left[\mathbb{E}_{(x,y)\sim\mathcal{D}}[(y_j - h(x)_j)\cdot b_i(h(x),a_i)] - \hat{\mathbb{E}}_D[(y_j - h(x)_j)\cdot b_i(h(x),a_i)]\right]\right| \\
&\le \mathbb{E}_{h\sim f}\left[\left|\mathbb{E}_{(x,y)\sim\mathcal{D}}[(y_j - h(x)_j)\cdot b_i(h(x),a_i)] - \hat{\mathbb{E}}_D[(y_j - h(x)_j)\cdot b_i(h(x),a_i)]\right|\right] \\
&\le \mathbb{E}_{h\sim f}\left[\sqrt{\frac{8\ln\frac{2|\mathcal{H}|Ndm}{\delta}}{n}}\right] = \sqrt{\frac{8\ln\frac{2|\mathcal{H}|Ndm}{\delta}}{n}}.
\end{aligned}$$

□

Now we are ready to prove Lemma 3.4.

Lemma 3.4. *We have*

$$|\mathcal{L}_{\mathcal{D}}(f,\lambda) - \mathcal{L}_D(f,\lambda)| \le \sqrt{\frac{\ln\frac{|4\mathcal{H}|}{\delta}}{2n}} + C\sqrt{\frac{8\ln\frac{4|\mathcal{H}|Ndm}{\delta}}{n}}$$

for all $f \in \Delta(\mathcal{H}), \lambda \in \Lambda$ with probability $1-\delta$.

Proof. This is straightforward from Theorem E.2 and Theorem E.3, we split the budget to $\delta/2$,

$$\begin{aligned}
&\mathcal{L}_D(f,\lambda) - \mathcal{L}_{\mathcal{D}}(f,\lambda) \\
=& -\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})] + -\mathbb{E}_{h\sim f}\mathbb{E}_D[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})] \\
&+ \sum_{s\in\{+,-\}}\sum_{i=1}^{N}\sum_{j=1}^{d}\sum_{a_i\in A_i}\lambda_{s,i,j,a_i}s(\mathbb{E}_f\mathbb{E}_{\mathcal{D}}[(h(x)_j - y_j)\cdot b_i(h(x),a_i)] - \gamma) \\
&- \sum_{s\in\{+,-\}}\sum_{i=1}^{N}\sum_{j=1}^{d}\sum_{a_i\in A_i}\lambda_{s,i,j,a_i}s(\mathbb{E}_f\mathbb{E}_{\mathcal{D}}[(h(x)_j - y_j)\cdot b_i(h(x),a_i)] - \gamma)
\end{aligned}$$

Therefore,

$$\begin{aligned}
&|\mathcal{L}_D(f,\lambda) - \mathcal{L}_{\mathcal{D}}(f,\lambda)| \\
\le&\left|\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})] - \mathbb{E}_{h\sim f}\mathbb{E}_D[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})]\right| + \sum_{s\in\{+,-\}}\sum_{i=1}^{N}\sum_{j=1}^{d}\sum_{a_i\in A_i} \\
&\lambda_{s,i,j,a_i}s|\mathbb{E}_f\mathbb{E}_{\mathcal{D}}[(h(x)_j - y_j)\cdot b_i(h(x),a_i)] - \mathbb{E}_f\mathbb{E}_D[(h(x)_j - y_j)\cdot b_i(h(x),a_i)]| \\
\le&\sqrt{\frac{\ln\frac{|4\mathcal{H}|}{\delta}}{2n}} + C\sqrt{\frac{8\ln\frac{4|\mathcal{H}|Ndm}{\delta}}{n}}.
\end{aligned}$$

□

Finally we are ready to prove Theorem 3.1.

Theorem 3.1. *Suppose* `PerDecCal` *runs for $T = O(\log(Nmd)/\epsilon^4)$ rounds and is given a dataset D drawn i.i.d. from $\mathcal{D}$ of size $n \ge O(\frac{\log(|\mathcal{H}|Ndm/\delta)}{\epsilon^4})$. With probability at least $1-\delta$, the output predictor $\hat{f}$ that satisfies*

1. $\mathrm{DecCE}(\hat{f}) \le \gamma + \epsilon$.
2. *Suppose the receivers play strict best response to $\hat{f}$. Then the receivers obtain swap regret bounded by $2mL(\gamma+\epsilon)$. The sender achieves ϵ-optimal utility:*

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}[u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})] \ge \mathrm{OPT}(\mathcal{H},\mathcal{D},\gamma) - \epsilon.$$

Proof. The regret bound of Hedge algorithm is $O(C\sqrt{T\log Nmd})$, and the best response of f give non-positive regret. From Lemma E.1 and $T = O(\log(Nmd/\delta)/\epsilon^4)$, we have that $(\hat{f}, \bar{\lambda})$ is an $\varepsilon/4$-approximate equilibrium under $\mathcal{L}_D(f, \lambda)$. From Lemma 3.4, we know that with probability $1-\delta$, $\forall f \in \Delta(\mathcal{H}), \lambda \in \Lambda$,

$$|\mathcal{L}_{\mathcal{D}}(f,\lambda) - \mathcal{L}_{\mathcal{D}}(f,\lambda)| \le \sqrt{\frac{\ln\frac{|4\mathcal{H}|}{\delta}}{2n}} + C\sqrt{\frac{8\ln\frac{4|\mathcal{H}|Ndm}{\delta}}{n}}.$$

Therefore, let $f' = \arg\min_{f\in\Delta(\mathcal{H})} \mathcal{L}_{\mathcal{D}}(f, \bar{\lambda})$

$$\begin{aligned}\mathcal{L}_{\mathcal{D}}(\hat{f},\bar{\lambda}) - \mathcal{L}_{\mathcal{D}}(f',\bar{\lambda}) &\le \mathcal{L}_{\mathcal{D}}(\hat{f},\bar{\lambda}) - \mathcal{L}_{D}(\hat{f},\bar{\lambda}) + \mathcal{L}_{D}(\hat{f},\bar{\lambda}) - \mathcal{L}_{D}(f',\bar{\lambda}) + \mathcal{L}_{D}(f',\bar{\lambda}) - \mathcal{L}_{\mathcal{D}}(f',\bar{\lambda}) \\ &\le \epsilon/2 + 2\sqrt{\frac{\ln\frac{|4\mathcal{H}|}{\delta}}{2n}} + 2C\sqrt{\frac{8\ln\frac{4|\mathcal{H}|Ndm}{\delta}}{n}}.\end{aligned}$$

Similarly, we can prove that

$$\max_{\lambda\in\Lambda} \mathcal{L}_{\mathcal{D}}(\hat{f},\lambda) - \mathcal{L}_{\mathcal{D}}(\hat{f},\bar{\lambda}) \le \epsilon/4 + 2\sqrt{\frac{\ln\frac{|4\mathcal{H}|}{\delta}}{2n}} + 2C\sqrt{\frac{8\ln\frac{4|\mathcal{H}|Ndm}{\delta}}{n}}.$$

Therefore, $(\hat{f}, \bar{\lambda})$ is an $\epsilon/4 + 2\sqrt{\frac{\ln\frac{|4\mathcal{H}|}{\delta}}{2n}} + 2C\sqrt{\frac{8\ln\frac{4|\mathcal{H}|Ndm}{\delta}}{n}}$-approximate equilibrium for the payoff $\mathcal{L}_{\mathcal{D}}(f, \lambda)$ under the true distribution $\mathcal{D}$. When $n \ge O(\frac{\log(|\mathcal{H}|Ndm)}{\epsilon^4})$, we have $(\hat{f}, \bar{\lambda})$ is a $\epsilon/2$-approximate equilibrium under the true distribution $\mathcal{D}$. Then, by Lemma 3.2, and the choice of C, we have

$$\mathrm{DecCE}(\hat{f}) \le \gamma + \frac{1+\epsilon}{C} \le \gamma + \frac{1+\epsilon}{2/\epsilon} \le \gamma + \frac{2}{2/\epsilon} = \gamma + \epsilon,$$

and

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}[u(\boldsymbol{a}, y)\cdot b(h(x), \boldsymbol{a})] \ge \mathrm{OPT}(\mathcal{H}, \mathcal{D}, \gamma) - \epsilon.$$

Finally by Theorem 2.5, the receivers who play η-quantal response obtain the stated swap regret bound. □

F Missing Proofs in Section 4

Lemma 4.2. *Consider a randomized predictor $f \in \Delta(\mathcal{H}_{\mathrm{ALL}})$ where $\mathcal{H}_{\mathrm{ALL}} = \{h : \mathcal{X} \to \mathcal{Y}\}$ is the class of all possible deterministic predictors. There exists a distribution $Q_f \in \Delta(\mathcal{Y})$ such that for any $v \in \mathcal{Y}$, $Q_f(v) = \Pr_{h\sim f,(x,y)\sim\mathcal{D}}[h(x) = v]$. A distribution $Q \in \Delta(\mathcal{Y})$ corresponds to the distribution over posterior means induced by some signaling scheme if and only if it is the prediction distribution Q_f of a perfectly calibrated predictor f.*

Proof. Fix any signaling scheme $\pi : \mathcal{Y} \to \Delta(S)$. Any context $x' \in \mathcal{X}$ corresponds to a state $\theta_{x'} = \mathbb{E}[y \mid x']$ and hence corresponds to a distribution of signals. Any signal $s \in S$ corresponds to a posterior mean $\mathbb{E}[\theta_{x'} \mid s]$. We define the mapping $g : \mathcal{X} \to \Delta(\mathcal{Y})$ such that

$$\Pr[g(x) = \mathbb{E}[\theta_{x'} \mid s]] = \pi(s \mid \theta_x).$$

We have that for any $s \in S$

$$\begin{aligned}\mathbb{E}[y \mid g(x) = \mathbb{E}[\theta_{x'} \mid s]] &= \int_y y \cdot \Pr[y \mid g(x) = \mathbb{E}[\theta_{x'} \mid s]]dy \\ &= \int_y y \cdot \left(\int_x \Pr[y \mid x, s]dx\right)dy \\ &= \int_x \left(\int_y \Pr[y \mid x, s]dy\right)dx \\ &= \int_x \mathbb{E}[y \mid x, s]dx \\ &= \int_x \theta_x \Pr[x \mid s]dx\end{aligned}$$

$$= \mathbb{E}[\theta_{x'} \mid s]$$

Therefore, we can convert g to a randomized predictor $f \in \Delta(\mathcal{H}_{\text{ALL}})$ that is perfectly calibrated.

Now consider any randomized predictor f that is calibrated, i.e. we have that for any $v \in \mathcal{Y}$

$$\mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}[y - h(x) \mid h(x) = v] = 0.$$

Now consider a signaling scheme such that the signal set is $\mathcal{Y}$ and given any state $\theta = \mathbb{E}[y \mid x]$, it sends signal $v \in \mathcal{Y}$ with probability $\Pr_{h\sim f}[h(x) = v \mid \theta]$. We have that for any signal v, the posterior mean is

$$\begin{aligned}\mathbb{E}[\theta \mid v] &= \mathbb{E}[\theta \mid h(x) = v] \\ &= \mathbb{E}[\mathbb{E}[y \mid x] \mid h(x) = v] \\ &= \mathbb{E}[y \mid h(x)] \\ &= v.\end{aligned}$$

□

Lemma 4.3. *For any randomized predictor f is perfectly decision calibrated, we can construct a randomized predictor f' such that (i) f' is perfectly calibrated; (ii) the sender obtains the same expected utility under f and f'.*

Proof. Without loss of generality, we assume that for any a, event $\{b(h(x), a) = 1\}$ happens with non-zero probability, otherwise we could remove that action. Since f is perfectly decision calibrated, we have

$$\max_{j\in[d]} \max_{a\in\mathcal{A}} \mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}[(y_j - h(x)_j) \cdot b(h(x), a)] = 0.$$

Since $\Pr\{b(h(x), a) = 1\} > 0$, equivalently, we have

$$\mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}[h(x) - y | b(h(x), a) = 1] = 0.$$

For any $a \in \mathcal{A}$, let $f'_a := \mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}[h(x)|b(h(x), a) = 1]$ Consider a post-processing function $p(h(x)) = \sum_{a\in\mathcal{A}} b(h(x), a)f'_a$. Now we are ready to construct f', we let $f' = p(f)$. Note that f' only output at most m values, we only need to check the level sets of each f'_a. Also, the set $\{y|b(y, a) = 1\}$ is convex, therefore $b(f'_a, a) = 1$.

$$\begin{aligned}&\mathbb{E}_{h\sim f'}\mathbb{E}_{(x,y)\sim\mathcal{D}}[(y_j - h(x)_j) \cdot \mathbf{1}(h(x) = f'_a)] \\ =&\mathbb{E}_{h\sim f'}\mathbb{E}_{(x,y)\sim\mathcal{D}}[(y_j - h(x)_j) \cdot b(h(x), a)] \\ =&\mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}[(y_j - p(h(x))_j) \cdot b(h(x), a)] \\ =&\mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}\left[\left(y_j - \left[\sum_{a'\in\mathcal{A}} b(h(x), a)f'_{a'}\right]_j\right) \cdot b(h(x), a)\right] \\ =&\mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}[(y_j - [b(h(x), a)f'_a]_j) \cdot b(h(x), a)] \\ =&\mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}[(y_j \cdot b(h(x), a)] - \mathbb{E}_{h\sim f}\mathbb{E}_{(x,y)\sim\mathcal{D}}[(h(x)_j \cdot b(h(x), a)] \\ =&0.\end{aligned}$$

□

Theorem 4.5. *Given at least $O(\frac{\ln(|\mathcal{H}|dm/\delta)}{\epsilon^4})$ samples, with probability $1 - \delta$, `PerDecCal` can output a predictor $\hat{f}$ such that the expected sender's utility under $\hat{f}$ is no worse than $\text{BayesOPT}(\mu_{\mathcal{D}}, \Pi_{\mathcal{H}}) - \epsilon$ where we denote the optimal sender utility under our Bayesian persuasion benchmark as $\text{BayesOPT}(\mu_{\mathcal{D}}, \Pi_{\mathcal{H}})$.*

Proof. By Theorem 3.1 we know that given any tolerance $\gamma \geq 0$ and $n \geq O(\frac{\ln(|\mathcal{H}|dm/\delta)}{\epsilon^4})$, let $C = \frac{2}{\epsilon}$, when `PerDecCal` runs for $T = O(\ln(Nmd)/\epsilon^4)$ rounds, with probability at least $1 - \delta$, it outputs $\hat{f}$ that satisfies

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}[u(\boldsymbol{a}, y) \cdot b(h(x), \boldsymbol{a})] \geq \text{OPT}(\mathcal{H}, \mathcal{D}, \gamma) - \epsilon$$

$$\geq \mathrm{OPT}(\mathcal{H}, \mathcal{D}, 0) - \epsilon.$$

Note that $\mathrm{OPT}(\mathcal{H}, \mathcal{D}, 0)$ is the optimal sender utility achieved by the randomized predictors over $\mathcal{H}$ that are perfectly decision calibrated, i.e. $\mathcal{F}_{\mathrm{DCAL}}(\mathcal{H})$. By Definition 4.4 and Lemma 4.3, $\mathrm{OPT}(\mathcal{H}, \mathcal{D}, 0)$ is equal to the optimal sender utility achieved by the predictors in $\mathcal{F}_{\mathrm{CAL}}(\mathcal{H})$.

Then by Definition 4.4 and Lemma 4.2, $\mathrm{OPT}(\mathcal{H}, \mathcal{D}, 0)$ is equal to the optimal sender utility achieved by the signaling schemes in $\Pi_{\mathcal{H}}$, i.e. $\mathrm{BayesOPT}(\mu_{\mathcal{D}}, \Pi_{\mathcal{H}})$.

Putting all these together, we know that the sender's expected utility under $\hat{f}$ is greater than $\mathrm{BayesOPT}(\mu_{\mathcal{D}}, \Pi_{\mathcal{H}}) - \epsilon$.

□

G Algorithm and Missing Proofs in Section 5

G.1 The algorithm for Infinite Hypothesis Class

Lagrangian and Minimax Game Similarly, we can introduce the Lagrangian and restrict the Lagrangian variables to be bounded as follows:

$$\min_{f \in \Delta\mathcal{H}} \max_{\lambda \in \Lambda} \tilde{\mathcal{L}}_{\mathcal{D}}(f, \lambda) := -\mathbb{E}_{h \sim f} \mathbb{E}_{\mathcal{D}}[\sum_{\boldsymbol{a} \in \mathcal{A}} u(\boldsymbol{a}, y) \cdot \tilde{b}(h(x), \boldsymbol{a})] + \sum_{s \in \{+,-\}} \sum_{i=1}^{N} \sum_{j=1}^{d} \sum_{a_i \in A_i} \lambda_{s,i,j,a_i} s\Big(\mathbb{E}_f \mathbb{E}_D[(h(x)_j - y_j) \cdot \tilde{b}_i(h(x), a_i)] - \gamma\Big). \tag{5}$$

Now we present our oracle efficient algorithm `SmPerDecCal`.

Algorithm 2 `SmPerDecCal`

Input: A set of samples D, ERM oracle $\mathrm{ERM}(D, \lambda)$, dual bound C and tolerance γ.

1: Initialize $\lambda_1 = \frac{C}{2Nmd}\mathbf{1}$.
2: **for** $t = 1, \cdots, T$ **do**
3: Learner best responds to λ_t:
4: Use the ERM oracle to compute $h_t = \mathrm{ERM}(D, \lambda_t)$.
5: Auditor runs Hedge to obtain λ_{t+1}:
6: $\lambda_{t+1} = \mathrm{Hedge}(c_{1:t})$ where $c_t(\lambda_{s,i,j,a_i}) = \lambda_{s,i,j,a_i} s(\mathbb{E}_D[(h_t(x) - y) \cdot \tilde{b}(h_t(x), a)] - \gamma)$.
7: **end for**

Output: Output $\hat{f} = \mathrm{Uniform}(h_1, \cdots, h_T)$.

G.2 Missing Proofs

We need the following technical lemmas before proving Theorem 5.4.

Lemma G.1 (Lipschitzness of single smoothed best response)**.** *For any $i \in [N], a_i \in \mathcal{A}_i$, we have that the function $\tilde{b}_i(\cdot, a_i)$ is $2\eta L$-Lipschitz in the L_∞ norm.*

Proof. For any $z \in [-1, 1]^d$, for simplicity, we drop the subscript i. Let $g_a(z) \triangleq \exp(\eta v(z, a))$, $G(z) \triangleq \sum_{a'} \exp(\eta v(z, a'))$. We have

$$\begin{aligned}
\nabla_z \tilde{b}(z, a) &= \frac{\nabla_z g_a(z)}{G(z)} - \frac{g_a(z) \nabla_z G(z)}{G(z)^2} \\
&= \frac{\eta g_a(z)}{G(z)} \nabla_z v(z, a) - \frac{\eta g_a(z)}{G(z)} \sum_{a'} \frac{g_{a'}(z)}{G(z)} \nabla_z v(z, a') \\
&= \eta b(z, a) \nabla_z v(z, a) - \sum_{a'} b(z, a') \nabla_z v(z, a').
\end{aligned}$$

Therefore for any $z \in [-1,1]^d$, we have

$$\begin{aligned}\left\|\nabla_z \tilde{b}(z,a)\right\|_1 &\le \eta\tilde{b}(z,a)\left(\|\nabla_z v(z,a)\|_1 + \sum_{a'}\tilde{b}(z,a')\|\nabla_z v(z,a')\|_1\right)\\ &\le \eta\tilde{b}(z,a)(L+L) \le 2\eta L.\end{aligned}$$

By mean-value theorem we have that

$$\begin{aligned}\left|\tilde{b}(z,a) - \tilde{b}(z',a)\right| &= \left|\int_0^1 \nabla_z \tilde{b}(z' + t(z-z'),a)\cdot(z-z')dt\right|\\ &\le \sup_{t\in[0,1]}\left\|\nabla_z\tilde{b}(z'+t(z-z'),a)\right\|_1 \|z-z'\|_\infty\\ &\le 2\eta L\|z-z'\|_\infty.\end{aligned}$$

□

Lemma G.2 (Lipschitzness of joint smoothed best response)**.** *For any $\boldsymbol{a} \in \mathcal{A}$, the joint smoothed best response $\tilde{b}(\cdot,\boldsymbol{a})$ satisfies that*

$$\sum_{\boldsymbol{a}\in\mathcal{A}}\left|\tilde{b}(z,\boldsymbol{a}) - \tilde{b}(z',\boldsymbol{a})\right| \le 2\eta m N L\|z-z'\|_\infty.$$

Proof. For any predictions $z, z' \in [-1,1]^d$ and joint action $\boldsymbol{a} = (a_1,\cdots,a_N)$, denote $b_i(z,a_i)$ as $u_i(a_i)$ and $b_i(z',a_i)$ as $v_i(a_i)$.

$$\begin{aligned}&\sum_{\boldsymbol{a}\in\mathcal{A}}|b(z,\boldsymbol{a}) - b(z',\boldsymbol{a})|\\ &= \sum_{\boldsymbol{a}\in\mathcal{A}}\left|\prod_{i\in[N]}u_i(a_i) - \prod_{i\in[N]}v_i(a_i)\right|\\ &= \sum_{\boldsymbol{a}\in\mathcal{A}}\left|\prod_{i\in[N]}u_i(a_i) - v_1(a_1)\prod_{i=2}^N u_i(a_i) + v_1(a_1)\prod_{i=2}^N u_i(a_i) - \prod_{i\in[N]}v_i(a_i)\right|\\ &= \sum_{\boldsymbol{a}\in\mathcal{A}}\left|(u_1(a_1)-v_1(a_1))\prod_{i=2}^N u_i(a_i) + v_1(a_1)\left(\prod_{i=2}^N u_i(a_i) - \prod_{i=2}^N v_i(a_i)\right)\right|\\ &= \sum_{\boldsymbol{a}\in\mathcal{A}}\left|(u_1(a_1)-v_1(a_1))\prod_{i=2}^N u_i(a_i) + v_1(a_1)(u_2(a_2)-v_2(a_2))\prod_{i=3}^N u_i(a_i)\right.\\ &\left. + v_1(a_1)v_2(a_2)\left(\prod_{i=3}^N u_i(a_i) - \prod_{i=3}^N v_i(a_i)\right)\right|\\ &\quad\vdots\\ &= \sum_{\boldsymbol{a}\in\mathcal{A}}\left|\sum_{i=1}^N\prod_{j=1}^{i-1}v_j(a_j)(u_i(a_i)-v_i(a_i))\prod_{j=i+1}^N u_j(a_j)\right|\\ &\overset{(a)}{\le} 2\eta L\|z-z'\|_\infty\sum_{\boldsymbol{a}\in\mathcal{A}}\left|\sum_{i=1}^N\prod_{j=1}^{i-1}v_j(a_j)\prod_{j=i+1}^N u_j(a_j)\right|\\ &= 2\eta L\|z-z'\|_\infty\sum_{\boldsymbol{a}\in\mathcal{A}}\sum_{i=1}^N\prod_{j=1}^{i-1}v_j(a_j)\prod_{j=i+1}^N u_j(a_j)\end{aligned}$$

$$
\begin{aligned}
&= 2\eta L\|z - z'\|_\infty \sum_{i=1}^{N} \sum_{\boldsymbol{a}\in\mathcal{A}} \prod_{j=1}^{i-1} v_j(a_j) \prod_{j=i+1}^{N} u_j(a_j) \\
&= 2\eta L\|z - z'\|_\infty \sum_{i=1}^{N} \sum_{a_i\in\mathcal{A}_i} \sum_{\boldsymbol{a}_{-i}} \prod_{j=1}^{i-1} v_j(a_j) \prod_{j=i+1}^{N} u_j(a_j) \\
&\overset{(b)}{=} 2\eta L\|z - z'\|_\infty \sum_{i=1}^{N} \sum_{a_i\in\mathcal{A}_i} 1 \\
&= 2\eta mNL\|z - z'\|_\infty
\end{aligned}
$$

where (a) holds because of Lemma G.1 and (b) holds because $\prod_{j=1}^{i-1} v_j(a_j) \prod_{j=i+1}^{N} u_j(a_j)$ can be viewed as the probability density function of $\boldsymbol{a}_{-i}$ and then $\sum_{\boldsymbol{a}_{-i}} \prod_{j=1}^{i-1} v_j(a_j) \prod_{j=i+1}^{N} u_j(a_j) = 1$. □

Lemma G.3. *Fix a class of deterministic predictors $\mathcal{H}$. For any distribution $\mathcal{D}$, let $D \sim \mathcal{D}^n$ be a dataset consisting of n samples $(x^{(i)}, y^{(i)})$ sampled i.i.d. from $\mathcal{D}$. Then for any $\delta \in (0,1)$, with $1-\delta$, for every $f \in \Delta(\mathcal{H})$, we have*

$$
\begin{aligned}
&\left| \mathbb{E}_{\mathcal{D}} \mathbb{E}_{h\sim f} \left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y) \cdot \tilde{b}(h(x), \boldsymbol{a}) \right] - \hat{\mathbb{E}}_D \mathbb{E}_{h\sim f} \left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y) \cdot \tilde{b}(h(x), \boldsymbol{a}) \right] \right| \\
&\leq \inf_{\epsilon>0} \left(4\eta mNL\epsilon + \sqrt{\frac{\ln \frac{2\mathcal{N}(\mathcal{H}, d_\infty, \epsilon)}{\delta}}{2n}} \right).
\end{aligned}
$$

Proof. Define $Z_h = |\mathbb{E}_{\mathcal{D}}[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y) \cdot \tilde{b}(h(x), \boldsymbol{a})] - \hat{\mathbb{E}}_D[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y) \cdot \tilde{b}(h(x), \boldsymbol{a})]|$. For any $h_1, h_2 \in \mathcal{H}$, we have

$$
\begin{aligned}
|Z_{h_1} - Z_{h_2}| = &\left| \mathbb{E}_{\mathcal{D}} \left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y) \cdot \tilde{b}(h_1(x), \boldsymbol{a}) - \tilde{b}(h_2(x), \boldsymbol{a})) \right] \right. \\
&\left. - \hat{\mathbb{E}}_D \left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y) \cdot (\tilde{b}(h_1(x), \boldsymbol{a}) - \tilde{b}(h_2(x), \boldsymbol{a})) \right] \right| \\
\overset{(a)}{\leq} &\left| \mathbb{E}_{\mathcal{D}} \left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y) \cdot (\tilde{b}(h_1(x), \boldsymbol{a}) - \tilde{b}(h_2(x), \boldsymbol{a})) \right] \right| \\
&+ \left| \hat{\mathbb{E}}_D \left[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a}, y) \cdot (\tilde{b}(h_1(x), \boldsymbol{a}) - \tilde{b}(h_2(x), \boldsymbol{a})) \right] \right| \\
\overset{(b)}{\leq} &\mathbb{E}_{\mathcal{D}} \left[\sum_{\boldsymbol{a}} \left| (\tilde{b}(h_1(x), \boldsymbol{a}) - \tilde{b}(h_2(x), \boldsymbol{a})) \right| \right] \\
&+ \hat{\mathbb{E}}_D \left[\sum_{\boldsymbol{a}} \left| (\tilde{b}(h_1(x), \boldsymbol{a}) - \tilde{b}(h_2(x), \boldsymbol{a})) \right| \right] \\
\overset{(c)}{\leq} &2\eta mNL(\mathbb{E}_{\mathcal{D}}[\|h_1(x) - h_2(x)\|_\infty] + \hat{\mathbb{E}}_D[\|h_1(x) - h_2(x)\|_\infty]) \\
\overset{(d)}{\leq} &4\eta mNL \sup_{x\in\mathcal{X}} \|h_1(x) - h_2(x)\|_\infty,
\end{aligned}
$$

where (a) holds because of Triangle Inequality, (b) holds because of $u(\boldsymbol{a}, y) \leq 1$, (c) holds because of Lemma G.2 and (d) holds by definition.

By Hoeffding's inequality, fixing any $h \in \mathcal{H}$, we have with probability $1-\delta$,

$$
|Z_h| \leq \sqrt{\frac{\ln \frac{2}{\delta}}{2n}}.
$$

Then as a result of the standard covering number argument we conclude with probability $1-\delta$ for any $h \in \mathcal{H}$,

$$|Z_h| \le 4\eta m N L\epsilon + \sqrt{\frac{\ln \frac{2\mathcal{N}(\mathcal{H}, d_\infty, \epsilon)}{\delta}}{2n}}$$

where $d_\infty(h_1, h_2) \triangleq \sup_{x\in\mathcal{X}} \|h_1(x) - h_2(x)\|_\infty$. □

Lemma G.4. *Fix a class of deterministic predictors $\mathcal{H}$. For any distribution $\mathcal{D}$, let $D \sim \mathcal{D}^n$ be a dataset consisting of n samples $(x^{(i)}, y^{(i)})$ sampled i.i.d. from $\mathcal{D}$. Then for any $\delta \in (0,1)$, with probability $1-\delta$, for every $f \in \Delta(\mathcal{H}), i \in [N], j \in [d], a_i \in \mathcal{A}_i$ we have*

$$\left|\mathbb{E}_{\mathcal{D}}\mathbb{E}_{h\sim f}\Big[(y_j - h(x)_j)\cdot \tilde{b}_i(h(x), a_i)\Big] - \hat{\mathbb{E}}_D\mathbb{E}_{h\sim f}\Big[(y_j - h(x)_j)\cdot \tilde{b}_i(h(x), a_i)\Big]\right|$$
$$\le \inf_{\epsilon>0}\left(8\eta L\epsilon + \sqrt{\frac{8\ln\frac{2\mathcal{N}(\mathcal{H}, d_\infty, \epsilon)Ndm}{\delta}}{n}}\right).$$

Proof. Define $Z_h = \mathbb{E}_{\mathcal{D}}[(y_j - h(x)_j)\cdot \tilde{b}_i(h(x), a_i)] - \hat{\mathbb{E}}_D[(y_j - h(x)_j)\cdot \tilde{b}_i(h(x), a_i)]$. For any $h_1, h_2 \in \mathcal{H}$, we have

$$\begin{aligned}
|Z_{h_1} - Z_{h_2}| = \Bigg|&\mathbb{E}_{\mathcal{D}}\Big[(y_j - h(x)_j)\cdot(\tilde{b}_i(h_1(x), a_i) - \tilde{b}_i(h_2(x), a_i))\Big] \\
&- \hat{\mathbb{E}}_D\Big[(y_j - h(x)_j)\cdot(\tilde{b}_i(h_1(x), a_i) - \tilde{b}_i(h_2(x), a_i))\Big]\Bigg| \\
\overset{(a)}{\le}& \left|\mathbb{E}_{\mathcal{D}}\Big[(y_j - h(x)_j)\cdot(\tilde{b}_i(h_1(x), a_i) - \tilde{b}_i(h_2(x), a_i))\Big]\right| \\
&+ \left|\hat{\mathbb{E}}_D\Big[(y_j - h(x)_j)\cdot(\tilde{b}_i(h_1(x), a_i) - \tilde{b}_i(h_2(x), a_i))\Big]\right| \\
\overset{(b)}{\le}& 2\left|\mathbb{E}_{\mathcal{D}}\Big[\tilde{b}_i(h_1(x), a_i) - \tilde{b}_i(h_2(x), a_i))\Big]\right| + 2\left|\hat{\mathbb{E}}_D\Big[\tilde{b}_i(h_1(x), a_i) - \tilde{b}_i(h_2(x), a_i))\Big]\right| \\
\overset{(c)}{\le}& 4\eta L\Big(\left|\mathbb{E}_{\mathcal{D}}[\|h_1(x) - h_2(x)\|_\infty]\right| + \left|\hat{\mathbb{E}}_D[\|h_1(x) - h_2(x)\|_\infty]\right|\Big) \\
\overset{(d)}{\le}& 8\eta L \sup_{x\in\mathcal{X}} \|h_1(x) - h_2(x)\|_\infty
\end{aligned}$$

where (a) holds because of Triangle Inequality, (b) holds because of $y, h(x) \in [-1,1]^d$, (c) holds because of Lemma G.1 and (d) holds by definition.

By Hoeffding's inequality, fixing and $h \in \mathcal{H}$, we have with probability $1-\delta$,

$$|Z_h| \le \sqrt{\frac{8\ln\frac{2}{\delta}}{n}}.$$

Then as a result of the standard covering number argument we conclude with probability $1-\delta'$ for any $h \in \mathcal{H}$,

$$|Z_h| \le 8\eta L\epsilon + \sqrt{\frac{8\ln\frac{2\mathcal{N}(\mathcal{H}, d_\infty, \epsilon)}{\delta'}}{n}}$$

where $d_\infty(h_1, h_2) \triangleq \sup_{x\in\mathcal{X}} \|h_1(x) - h_2(x)\|_\infty$.

Then let $\delta' = \delta/(Ndm)$, by union bound we have with probability $1-\delta$, for any $h \in \mathcal{H}$, $i \in [N]$ and $j \in [d]$ and $a_i \in \mathcal{A}_i$,

$$\mathbb{E}_{\mathcal{D}}\Big[(y_j - h(x)_j)\cdot \tilde{b}_i(h(x), a_i)\Big] - \hat{\mathbb{E}}_D\Big[(y_j - h(x)_j)\cdot \tilde{b}_i(h(x), a_i)\Big] \le 8\eta L\epsilon + \sqrt{\frac{8\ln\frac{2\mathcal{N}(\mathcal{H}, d_\infty, \epsilon)Ndm}{\delta}}{n}}.$$

□

Lemma G.5. *With probability* $1-\delta$, $\forall f \in \Delta(\mathcal{H}), \lambda \in \Lambda$,

$$\begin{aligned}&\left|\tilde{\mathcal{L}}_D(f,\lambda) - \tilde{\mathcal{L}}_{\mathcal{D}}(f,\lambda)\right| \\ \leq &\inf_{\epsilon>0}\left(4\eta(mN+2C)L\epsilon + \sqrt{\frac{\ln\frac{4\mathcal{N}(\mathcal{H},d_\infty,\epsilon)}{\delta}}{2n}} + C\sqrt{\frac{8\ln\frac{4\mathcal{N}(\mathcal{H},d_\infty,\epsilon)Ndm}{\delta}}{n}}\right).\end{aligned}$$

Proof. This is straightforward from Lemma G.3 and Lemma G.4. First we have

$$\begin{aligned}&\tilde{\mathcal{L}}_D(f,\lambda) - \tilde{\mathcal{L}}_{\mathcal{D}}(f,\lambda) \\ =& -\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot\tilde{b}(h(x),\boldsymbol{a})] + -\mathbb{E}_{h\sim f}\mathbb{E}_D[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot\tilde{b}(h(x),\boldsymbol{a})] \\ &+ \sum_{s\in\{+,-\}}\sum_{i=1}^{N}\sum_{j=1}^{d}\sum_{a_i\in A_i}\lambda_{s,i,j,a_i}s\Big(\mathbb{E}_f\mathbb{E}_{\mathcal{D}}[(h(x)_j - y_j)\cdot\tilde{b}_i(h(x),a_i)] - \gamma\Big) \\ &- \sum_{s\in\{+,-\}}\sum_{i=1}^{N}\sum_{j=1}^{d}\sum_{a_i\in A_i}\lambda_{s,i,j,a_i}s\Big(\mathbb{E}_f\mathbb{E}_{\mathcal{D}}[(h(x)_j - y_j)\cdot\tilde{b}_i(h(x),a_i)] - \gamma\Big)\end{aligned}$$

Therefore, by union bound with probability $1-\delta$,

$$\begin{aligned}&\left|\tilde{\mathcal{L}}_D(f,\lambda) - \tilde{\mathcal{L}}_{\mathcal{D}}(f,\lambda)\right| \\ \leq&\left|\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})] - \mathbb{E}_{h\sim f}\mathbb{E}_D[\sum_{\boldsymbol{a}\in\mathcal{A}} u(\boldsymbol{a},y)\cdot b(h(x),\boldsymbol{a})]\right| + \sum_{s\in\{+,-\}}\sum_{i=1}^{N}\sum_{j=1}^{d}\sum_{a_i\in A_i} \\ &\lambda_{s,i,j,a_i}s|\mathbb{E}_f\mathbb{E}_{\mathcal{D}}[(h(x)_j - y_j)\cdot b_i(h(x),a_i)] - \mathbb{E}_f\mathbb{E}_D[(h(x)_j - y_j)\cdot b_i(h(x),a_i)]| \\ \leq&\inf_{\epsilon>0}\left(4\eta(mN+2C)L\epsilon + \sqrt{\frac{\ln\frac{4\mathcal{N}(\mathcal{H},d_\infty,\epsilon)}{\delta}}{2n}} + C\sqrt{\frac{8\ln\frac{4\mathcal{N}(\mathcal{H},d_\infty,\epsilon)Ndm}{\delta}}{n}}\right).\end{aligned}$$

□

Now we are ready to prove Theorem 5.4.

Theorem 5.4. *Suppose* `SmPerDecCal` *runs for* $T = O(\log(Nmd)/\epsilon^4)$ *rounds and is given a dataset* D *drawn i.i.d of size* $n \geq O(\ln\frac{\mathcal{N}(\mathcal{H},d_\infty,\frac{\epsilon^2}{\eta L})Ndm}{\delta}/\epsilon^4)$. *With probability at least* $1-\delta$, *it outputs* $\hat{f}$ *that satisfies*

1. $\mathrm{SmDecCE}(\hat{f}) \leq \gamma + \epsilon$.
2. *Suppose the receivers play* η*-quantal response to* $\hat{f}$. *Then the receivers obtain swap regret bounded by* $2mL(\gamma+\epsilon) + \frac{\ln m+1}{\eta}$. *The sender achieves* ϵ*-optimal utility:*

$$\mathbb{E}_{h\sim f}\mathbb{E}_{\mathcal{D}}\Big[u(\boldsymbol{a},y)\cdot\tilde{b}(h(x),\boldsymbol{a})\Big] \geq \mathrm{OPT}(\mathcal{H},\mathcal{D},\gamma) - \epsilon.$$

Proof. The regret bound of the Hedge algorithm is $O(C\sqrt{T\log Nmd})$, and the ERM oracle gives non-positive regret. From Lemma E.1 and $T = O(\log(Nmd)/\epsilon^4)$, we have that $(\hat{f},\bar{\lambda})$ is an $\varepsilon/4$-approximate equilibrium under $\tilde{\mathcal{L}}_D(f,\lambda)$. From Lemma G.5, we know that when $n > \alpha(\epsilon,\delta,\eta,m,d,N,L) \triangleq O(\ln\frac{\mathcal{N}(\mathcal{H},d_\infty,\frac{\epsilon^2}{\eta L})Ndm}{\delta}/\epsilon^4)$ with probability $1-\delta$, $\forall f\in\Delta(\mathcal{H}), \lambda\in\Lambda$,

$$\left|\tilde{\mathcal{L}}_{\mathcal{D}}(f,\lambda) - \tilde{\mathcal{L}}_{\mathcal{D}}(f,\lambda)\right| \leq \frac{\epsilon}{8}.$$

Therefore, let $f' = \arg\min_{f\in\Delta(\mathcal{H})}\mathcal{L}_{\mathcal{D}}(f,\bar{\lambda})$

$$\tilde{\mathcal{L}}_{\mathcal{D}}(\hat{f},\bar{\lambda}) - \tilde{\mathcal{L}}_{\mathcal{D}}(f',\bar{\lambda}) \leq \tilde{\mathcal{L}}_{\mathcal{D}}(\hat{f},\bar{\lambda}) - \tilde{\mathcal{L}}_D(\hat{f},\bar{\lambda})$$

Table 1: HELOC test-set results (single receiver), averaged over three random seeds.

Method	Sender utility ↑	DC error ↓
ERM (cross-entropy)	0.026 ± 0.003	0.007 ± 0.004
`SmPerDecCal`	0.163 ± 0.006	0.008 ± 0.001

Table 2: HELOC test-set results (three receivers), averaged over three random seeds.

Method	Sender utility ↑	DC error ↓
ERM (cross-entropy)	0.060 ± 0.004	0.007 ± 0.004
`SmPerDecCal`	0.193 ± 0.005	0.011 ± 0.001

$$
\begin{aligned}
&+ \tilde{\mathcal{L}}_D(\hat{f}, \bar{\lambda}) - \tilde{\mathcal{L}}_D(f', \bar{\lambda}) + \tilde{\mathcal{L}}_D(f', \bar{\lambda}) - \tilde{\mathcal{L}}_{\mathcal{D}}(f', \bar{\lambda}) \\
\leq &\frac{\epsilon}{8} + \frac{\epsilon}{4} + \frac{\epsilon}{8} = \frac{\epsilon}{2}.
\end{aligned}
$$

Similarly, we can prove that

$$
\max_{\lambda \in \Lambda} \tilde{\mathcal{L}}_{\mathcal{D}}(\hat{f}, \lambda) - \tilde{\mathcal{L}}_{\mathcal{D}}(\hat{f}, \bar{\lambda}) \leq \frac{\epsilon}{2}.
$$

Therefore, $(\hat{f}, \bar{\lambda})$ is an $\epsilon/2$-approximate equilibrium for the payoff $\mathcal{L}_{\mathcal{D}}(f, \lambda)$ under the true distribution $\mathcal{D}$. Then, similar to Lemma 3.2, we have

$$
\mathrm{DecCE}(\hat{f}) \leq \gamma + \frac{1+\epsilon}{C} \leq \gamma + \frac{1+\epsilon}{2/\epsilon} \leq \gamma + \frac{2}{2/\epsilon} = \gamma + \epsilon,
$$

and

$$
\mathbb{E}_{h \sim f} \mathbb{E}_{\mathcal{D}}[u(\boldsymbol{a}, y) \cdot \tilde{b}(h(x), \boldsymbol{a})] \geq \mathrm{OPT}(\mathcal{H}, \mathcal{D}, \gamma) - \epsilon.
$$

Finally by Theorem C.8, the receivers who play η-quantal response obtain the stated swap regret bound. □

H Empirical validation on a real-world dataset

To better connect our setting to a realistic lending application, we run experiments on the public FICO HELOC dataset [FICO, 2018], which contains de-identified home-equity line of credit applications together with a binary `RiskPerformance` label indicating whether the borrower eventually became seriously delinquent. We model a stylized lending scenario in which a predictor (the sender) provides a probability-of-default forecast to a bank (the receiver), which then decides whether to approve or deny the loan.

H.1 Methods

We compare standard cross-entropy empirical risk minimization (ERM) with `SmPerDecCal`, using the same multilayer perceptron (MLP) function class for both methods. Following our theoretical framework, `SmPerDecCal` enforces smoothed decision calibration through the Lagrangian formulation defined in Eq. (3) (replacing b to $\tilde{b}$). In practice, we additionally include a log-loss constraint implemented via an extra dual variable, which empirically improves optimization stability and preserves the prediction accuracy.

H.2 Single-receiver setting

Setup. We first consider a single-receiver instantiation that directly corresponds to the motivating lending example. This setting captures a natural incentive misalignment: the bank is conservative and prefers to approve a loan only when the predicted default risk is sufficiently low, reflecting the asymmetric cost of defaults, whereas the predictor prefers approvals whenever defaults are unlikely.

We encode preferences using linear utilities over actions and outcomes. Denials yield utility 0 for both parties. For approvals, utilities are affine functions of the outcome $y \in \{0, 1\}$ (where $y = 1$ denotes default). The sender's approval utility is

$$u(\texttt{approve}, y) = 2.5 - 3y,$$

while the receiver's approval utility is

$$v(\texttt{approve}, y) = 0.1 - 2y.$$

Results. Table 1 reports test-set performance in the single-receiver setting, averaged over three random seeds. `SmPerDecCal` consistently achieves higher sender utility than ERM while maintaining comparable smoothed decision-calibration error, showing that decision calibration can improve sender outcomes even in this simple setting.

H.3 Multi-receiver setting

Setup. We next evaluate a three-receiver variant with heterogeneous receiver preferences, which introduces a stronger form of incentive misalignment. The sender's utility is additive across receivers, using the same per-receiver sender utility as in the single-receiver case and aggregating with fixed weights $(0.2, 0.3, 0.5)$. Each receiver i has an approval utility of the form $v_i(\texttt{approve}, y) = a_i y + b_i$, with

$$v_1(\texttt{approve}, y) = -3.0y + 0.30, \quad v_2(\texttt{approve}, y) = -2.0y + 0.10, \quad v_3(\texttt{approve}, y) = -1.5y + 0.12,$$

and $v_i(\texttt{deny}, y) = 0$ for all i. These utilities correspond to receivers with different degrees of conservatism.

Results. Table 2 summarizes performance in the multi-receiver setting. Compared to ERM, `SmPerDecCal` yields a substantially larger improvement in sender utility than in the single-receiver case, while controlling the smoothed decision-calibration error to be 0.011. Empirically, we notice the performance gap between `SmPerDecCal` and ERM depends on the choice of utility parameters: stronger misalignment between sender and receivers leads to a larger improvement in sender utility.

I Numerical validation of Theorem 4.5

We provide a finite-hypothesis numerical validation of Theorem 4.5. The purpose of this experiment is to construct a small instance where the restricted Bayesian benchmark $\mathrm{BayesOPT}(\mu_{\mathcal{D}}, \Pi_{\mathcal{H}})$ can be computed exactly and then compare it with the utility achieved by `PerDecCal`.

Setup. We consider a binary feature space $\mathcal{X} = \{0, 1\}$ with

$$\Pr[X = 0] = \Pr[X = 1] = 0.5, \qquad \mathbb{E}[Y \mid X = 0] = 0.2, \quad \mathbb{E}[Y \mid X = 1] = 0.8.$$

The finite hypothesis class is

$$\mathcal{H} = \{0.1, 0.3, 0.7, 0.9\}^2,$$

so each $h \in \mathcal{H}$ assigns one prediction to each feature value and $|\mathcal{H}| = 16$. The receiver follows a strict best response, approving iff $h(x) > 0.5$. The sender utility is

$$u(\texttt{reject}, y) = 0, \qquad u(\texttt{approve}, y) = 0.25 + 0.75y.$$

Since $\mathcal{X}$ and $\mathcal{H}$ are finite, we compute $\mathrm{BayesOPT}(\mu_{\mathcal{D}}, \Pi_{\mathcal{H}})$ exactly by solving a linear program over randomized predictors $f \in \Delta(\mathcal{H})$ that satisfy perfect decision calibration. We then run sample-based `PerDecCal` with target calibration level $\gamma = 0.05$, using $n = 1024$ samples, and evaluate the resulting randomized predictor on the true distribution $\mathcal{D}$.

Results. Table 3 reports the exact restricted Bayesian benchmark and the finite-sample performance of `PerDecCal`. The exact benchmark is

$$\mathrm{BayesOPT}(\mu_{\mathcal{D}}, \Pi_{\mathcal{H}}) = 0.465.$$

With $n = 1024$ samples, `PerDecCal` achieves average sender utility 0.504 ± 0.001 and average decision-calibration error 0.049 ± 0.002 over 50 random seeds, where error bars denote standard

Table 3: Finite-hypothesis validation of Theorem 4.5. The sample-based `PerDecCal` result uses $n = 1024$ samples and is averaged over 50 random seeds; error bars denote standard errors.

Method/Benchmark	Sender utility ↑	DC error ↓	Success rate ↑
$\mathrm{BayesOPT}(\mu_{\mathcal{D}}, \Pi_{\mathcal{H}})$	0.465	0	–
`PerDecCal`, $\gamma = 0.05$	0.504 ± 0.001	0.049 ± 0.002	0.940 ± 0.034

errors. We also report a success rate of 0.940 ± 0.034, where a run is counted as successful if its decision-calibration error is at most $\gamma + 0.02$ and its sender utility is at least $0.505 - 0.02$, where 0.505 is the best utility achieved by any mixture over $\mathcal{H}$ with decision-calibration error at most $\gamma = 0.05$. Thus, the learned randomized predictor satisfies the target decision-calibration level and achieves sender utility above the stricter perfectly decision-calibrated Bayesian benchmark, consistent with Theorem 4.5.

J Experiments compute resources

The finite-hypothesis validation runs on CPU and takes under one minute. The HELOC experiments use small MLP models on a single NVIDIA A100 GPU with 80GB of memory; the reported single-receiver and three-receiver three-seed runs together take less than 30 minutes in total. Memory and storage requirements are modest, since the HELOC dataset is a small tabular CSV file.

NeurIPS Paper Checklist

1. **Claims**

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

   Answer: [Yes]

   Justification: We clearly state our result on persuasive prediction and explain its connection to standard Bayesian persuasion.

   Guidelines:

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

2. **Limitations**

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

   Answer: [Yes]

   Justification: We discuss the restricted Bayesian benchmark in Section 4 and the discontinuity of strict best responses in Section 5.

   Guidelines:

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

3. **Theory assumptions and proofs**

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

Answer: [Yes]

Justification: The assumptions are clearly stated, and the full proofs are in the appendix.

Guidelines:

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

4. **Experimental result reproducibility**

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

   Answer: [Yes]

   Justification: We report necessary information in Appendix H and Appendix I.

   Guidelines:

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

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

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

Answer: [No]

Justification: Code will be released upon paper acceptance.

Guidelines:

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

6. **Experimental setting/details**

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

Answer: [Yes]

Justification: We report necessary information in Appendix H and Appendix I.

Guidelines:

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

7. **Experiment statistical significance**

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

Answer: [Yes]

Justification: Error bars are provided for all our experiments.

Guidelines:

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

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

8. **Experiments compute resources**

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

   Answer: [Yes]

   Justification: The experiments are lightweight. The finite-hypothesis validation runs on CPU and takes under one minute. The HELOC experiments use small MLP models on a single NVIDIA GPU; the reported single-receiver and three-receiver three-seed runs together take less than 30 minutes in total. Memory and storage requirements are modest, since the HELOC dataset is a small tabular CSV file. We include the above information in Appendix J.

   Guidelines:

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

9. **Code of ethics**

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

   Answer: [Yes]

   Justification: This paper conforms, in every respect, with the NeurIPS Code of Ethics.

   Guidelines:

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

10. **Broader impacts**

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

    Answer: [N/A]

    Justification: This paper is mainly theoretical and does not have any direct societal impact.

    Guidelines:

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

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

11. **Safeguards**

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

    Answer: [N/A]

    Justification: This paper is a theory paper and therefore does not pose such risks.

    Guidelines:

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

12. **Licenses for existing assets**

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

    Answer: [Yes]

    Justification: All datasets utilized in this study are appropriately cited and credited within the manuscript.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.

- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

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

    Answer: [Yes]

    Justification: Code will be released upon paper acceptance.

    Guidelines:

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

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

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

    Answer: [N/A]

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

    Guidelines:

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

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

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

    Answer: [N/A]

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

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.

- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

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

    Answer: [N/A]

    Justification: LLM is used only for writing, editing, or formatting purposes.

    Guidelines:

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