

Algorithmic Information Design for Searchers with Uncertain Alternatives¹

Zhicheng Du², Hu Fu³, Ying Qin⁴, and Ziheng Wang⁵

Abstract. Advertisements reveal information to consumers who decide on further information acquisition and eventual purchase. [Anderson and Renault \(2006\)](#) first modeled this problem using an information-design framework where the advertiser acts as a sender and the consumer as a receiver. Due to search frictions and the consumer’s outside option, search for additional information is not always worthwhile for the receiver. Thus, the sender’s information design is used to make further search attractive and ultimately induce purchase. Following [Lyu \(2023\)](#), we study the optimal information design for a sender who sells search goods to a consumer with uncertain alternatives. Instead of making relaxations to the sender’s problem (as done in [Lyu, 2023](#)), we work directly on the joint distribution over realized values and signals. Our contributions are twofold. First, we give a method, based on duality arguments, to verify whether a given information strategy is optimal. We illustrate the value of this verification framework in a competitive extension, where it certifies a non-trivial symmetric equilibrium for two senders with a common convex prior. Second, on the algorithmic front, we develop an FPTAS that finds for the seller a signaling scheme whose utility differs from that of the optimal solution by at most an additive ϵ error, for any $\epsilon > 0$.

Keywords: Information Design · Sequential Search · Optimality Verification · FPTAS

1 Introduction

Consumers are often faced with horizontally differentiated products without being fully informed of the attributes offered by different sellers. This information gap is often bridged, at least partially, by the advertisements posted by sellers to compete for the purchase. This paper takes an information design perspective to study the design of informative advertising for buyers engaged in a sequential search.

In [Anderson and Renault \(2006\)](#)’s influential model of *informative advertising*, an advertisement is viewed as a signal sent by the seller to directly, though perhaps only partially, inform a potential buyer about her valuation of a product. The signal, bound by legal sanctions, is credible. The unit-demand buyer, with posterior estimates from the advertisements, searches for a purchase.

In this context, advertisements are only meaningful when there is search friction. [Anderson and Renault](#) assumed the buyer incurs an inspection cost in order to exactly learn her value for any product; they study a monopoly seller facing a buyer with a deterministic outside option. Following [Lyu \(2023\)](#), we study an extended framework which allows the buyer to have a randomized value for her outside option. This extension, as we explain in Section 2.2, makes the framework applicable to a seller in a competitive market for search goods, with the outside option summarizing the competition from all rival sellers. This single-dimensional summarization is made possible by structural properties of the buyer’s optimal search strategy, the celebrated *Index Algorithm* first discovered by [Weitzman \(1978\)](#) for the *Pandora Box* problem.

[Anderson and Renault](#) studied a seller that strategically sets a price and commits to an advertising signaling scheme. As in [Lyu \(2023\)](#), we abstract away the pricing part and only consider strategic signaling: a seller gets utility 1 when the buyer makes a purchase from him,

¹ An earlier version of this paper was titled *Competitive Information Design in Sequential Search*.

² Renmin University of China. Email: duzhicheng@ruc.edu.cn

³ Shanghai University of Finance and Economics. Email: fu.hu.thu@gmail.com

⁴ Renmin University of China. Email: qinying0420@ruc.edu.cn

⁵ Renmin University of China. Email: wang.zihe@ruc.edu.cn

and 0 otherwise.¹ Unlike the case with deterministic outside option, the information design part in our setting is already very complicated. This simplification is also meaningful because decisions on pricing and promotion/marketing are often made by different departments in modern corporations; advertising and pricing may be decoupled decisions (It is for similar reasons that bidders in ad auctions are sometimes modeled as “value maximizers” rather than standard utility maximizers; see, e.g., Wilkens et al., 2016, 2017; Deng et al., 2024; Balseiro et al., 2024.)

We study the seller’s strategic advertising as an instance of information design (Kamenica and Gentzkow, 2011). A unit-demand buyer has match value v for the seller’s good; v , drawn from a publicly known distribution F , is unknown to the buyer initially, but can be learned if the buyer inspects it at a cost $c > 0$. The buyer’s value o for her outside option is drawn from a publicly known distribution H , but o is private to the buyer. The seller aims to maximize the probability that the buyer searches for and selects his item. To this end, the sender, prior to any inspection, commits to a signaling scheme that stochastically maps v to a signal. The signal acts as an advertisement that updates the buyer’s belief about v to a posterior distribution. The buyer implements the Index Algorithm to decide whether to inspect and whether to take the seller’s good. In the language of information design, the buyer is the *receiver*, and the seller the *sender*.

Our Contributions. We make contributions on two fronts of the sender’s persuasion problem. First, we develop a procedure for verifying the optimality of the seller’s signaling scheme and apply it to identify equilibria in markets with two advertising sellers. This is the first time such equilibria are characterized in an oligopoly; while Lyu (2023) was only able to identify equilibria when the number of sellers goes to infinity (the large-market limit). On the algorithmic front, we develop a fully polynomial-time approximation scheme (FPTAS) which finds for the seller a signaling scheme whose yield utility differs from that of the optimal solution by at most an additive ϵ error, for any $\epsilon > 0$. Underlying both results is a lossless representation of the sender’s signaling as a value-index joint distribution.

As we explain in Section 2, this two-dimensional representation naturally results from the Index Algorithm. Working with it directly poses technical difficulty: the feasible set of strategies is given by an infinite-dimensional linear program with a continuum of feasibility constraints. Lyu avoided the difficulty by representing the signaling scheme as a *one-dimensional* distribution of the so-called amortized values (see Definition 3); they imposed on these distributions a relaxed, necessary feasibility condition. Whereas this relaxation allows the use of tools from the persuasion literature (such as the techniques worked on mean-preserving contractions), it comes with loss. Generally, solutions to this relaxed program may not be feasible. Lyu derived conditions under which the relaxed program yields a feasible optimal solution; these conditions include quasi-concavity (and hence unimodality) of H , the distribution of the outside option value. Notably, when one applies the framework to study a competitive market with multiple sellers (which motivates the single-seller setting in the first place), a constraint on H in the single seller setting does not translate into a constraint on the value distribution F , the only input in the market setting. This is because H is used to summarize rival sellers’ competition; in an equilibrium, it incorporates information about the rival sellers’ endogenous signaling. Therefore, in a symmetric equilibrium, an assumption on H is really an assumption on the *solution* of the program. It is for this reason that Lyu was able to derive equilibria only when the number of sellers goes to infinity.

We tackle the formidable two-dimensional distribution head-on. Notice first that our approach imposes neither relaxation nor restriction on the original problem. On one front, we exploit the geometry of the value-index joint distributions and develop dual certificates for the

¹ Strictly speaking, the last part of Lyu (2023) discusses pricing, but in essence does not go beyond saying that the seller’s revenue is the product of his price and his demand (which is our utility, inherently a function of the price), and should be optimized by selecting the best price. The mathematical substance of the work almost entirely lies in the part that considers an exogenous price.

optimality of signaling schemes (Theorem 1). The value of this verification framework goes beyond equilibrium certification, although it is one important application. It is worth pointing out that the two-seller equilibrium we derive (Theorem 2) does not satisfy Lyu’s unimodality assumption. On the other front, we develop an additive FPTAS (Theorem 3) which uses a two-level discretization that performs the two dimensions, realized values and posterior indices, in different ways. We also study the competitive extension with multiple senders in several additional directions, including equilibrium existence in the general game and comparisons of information disclosure across different competitive settings. Due to space constraints, these results are deferred to Appendices D to F.

1.1 Related Literature

Our work contributes to two related strands of literature.

Sequential search and Pandora Box. In our model, the receiver faces an uncertain outside option and a search good controlled by the sender. She decides whether to pay the cost to learn the item’s value and then choose the better option, or to take the outside option directly. This sequential decision problem can be seen as an instance of the *Pandora Box* problem introduced in Weitzman (1978). The Pandora Box problem has motivated a rapidly growing algorithmic literature on sequential search and its variants (e.g. Doval, 2018; Beyhaghi and Kleinberg, 2019; Boodaghians et al., 2020; Chawla et al., 2020; Fu et al., 2023; Beyhaghi and Cai, 2023a; Hajiaghayi et al., 2025; Banihashem et al., 2025). We refer to Beyhaghi and Cai (2023b) for a survey. A central insight of Weitzman (1978) is that the receiver’s optimal search policy is governed by index values. The Index Algorithm has also been widely used to model consumer search in markets with competing sellers. See, for example, Armstrong (2017) for a survey, and Derakhshan et al. (2022) and Friedler et al. (2025) for recent applications. Our work differs from this literature by focusing on the sender’s optimal pre-search information design.

Information design for search goods. We model the senders’ pre-search information revelation as an instance of Bayesian persuasion (Kamenica and Gentzkow, 2011). This framework has inspired a large body of works on information design. See, e.g., Bergemann and Morris, 2019; Dughmi, 2017; Kamenica, 2019 for some recent surveys.

A large strand of the literature studies information provision for experience goods, where the receiver cannot learn the exact value through inspection before purchase and must make the purchase decision based only on the information provided by the seller. More recent work studies information disclosure and advertising in monopoly and competitive markets, including Ivanov (2013); Boleslavsky et al. (2017); Hwang et al. (2026). Related persuasion tools include the concavification approach of Kamenica and Gentzkow (2011), the linear-programming approach of Kolotilin (2018), and the duality-based approaches of Dworczak and Martini (2019); Dworczak and Kolotilin (2024).

However, information design for search goods is less developed. The main obstacle is that what the sender controls is the receiver’s pre-search information. Thus, posterior means alone are no longer sufficient for determining outcomes due to the inherently complicated search decision. Anderson and Renault (2006) initiate this line by studying advertising search goods to consumers with deterministic alternatives. Anderson and Renault (2013) extend the framework to allow for richer advertising content. Several related papers focus on post-search information design instead of pre-search information design, including Bar-Isaac et al. (2012); Board and Lu (2018); Whitmeyer (2020); Dogan and Hu (2022); Au and Whitmeyer (2023). Closer to pre-search information design for search goods, Choi et al. (2019) characterize consumer-optimal and seller-worst signals in a binary-value model, while Hinnosaar and Kawai (2020) study seller-worst pre-search information in a robust-pricing environment. In contrast to these works, we study the seller’s optimal pre-search information strategy with a continuous value distribution and a buyer with a privately known outside option.

The closest paper to ours is [Lyu \(2023\)](#), who first studies optimal pre-search information provision by a search-goods seller with privately known outside options. He reduces the sender’s persuasion problem to a one-dimensional relaxed problem over the distribution of the amortized values, and shows that the relaxed optimum solves the original problem only under unimodality conditions on the outside-option distribution. Our paper complements this work by keeping the exact value-index formulation without making any relaxation and developing algorithmic tools for the original best-response problem.

Roadmap. Section 2 introduces the model and identifies the sender’s persuasion problem. Section 3 develops the dual-certificate framework for verifying optimality of candidate strategies. Section 4 provides an additive FPTAS for computing near-optimal information strategies.

2 Preliminaries

2.1 The Setup

We consider a single strategic sender (he) who faces a receiver (she) with uncertain alternatives. The sender controls one item whose match value to the receiver is $v \in [0, 1]$. The value is drawn from a commonly known prior $F \in \Delta([0, 1])$,² and is *initially invisible* to the receiver.

The receiver can pay an inspection cost $c > 0$ to learn the realized value v (i.e., the item is a search good). The receiver also has an outside option $o \in [0, 1]$, drawn from a commonly known distribution $H \in \Delta([0, 1])$. The realized outside option is the receiver’s private information.

Search friction makes the additional information generated by inspection not always worthwhile. The seller can design a pre-search signal about the realized value v to attract the receiver. Following the information design literature, the sender commits to a signaling scheme $(\mathcal{S}, \pi)$, where $\mathcal{S}$ is the signal space and $\pi(\cdot | v)$ specifies the conditional distribution of signals given the realized value v .

The timing of the game is as follows.

1. The sender publicly commits to a signaling scheme $(\mathcal{S}, \pi)$.
2. The sender observes value $v \sim F$, and sends signal $s \sim \pi(\cdot | v)$ to the receiver.
3. The receiver updates the posterior belief about v after receiving signal s .
4. The receiver observes outside option $o \sim H$, and decides whether to search for the item.
 - (4.1) If not, she takes the outside option o and leaves the game.
 - (4.2) If yes, she incurs a search cost $c > 0$ and learns v exactly. Then she takes the larger value among v and o and leaves the game. Ties are broken uniformly at random.

2.2 The Receiver’s Optimal Search

We review [Weitzman \(1978\)](#)’s Pandora Box problem which models a sequential searcher’s problem and admits an elegant, index-based optimal strategy. Our purpose is to illustrate (a) that the setup in Section 2.1 is an abstraction of a sender’s problem in a search good market with multiple competing sellers, and (b) that the outcome of the optimal search is completely determined by a comparison between the outside option and the sender’s so-called amortized values.

In the Pandora Box problem, a searcher is faced with n boxes, each with a value that becomes observable only after an inspection cost is paid. Specifically, each box i has value v_i drawn independently from a publicly known distribution F_i , and is revealed at inspection cost c_i . The searcher can take at most one box, and must optimize her search order and stopping time to maximize the expected value of the box taken minus the cumulated inspection costs. [Weitzman](#) showed the following *Index Algorithm* to be the optimal strategy.

² Throughout the paper, $\Delta(\cdot)$ denotes the set of all probability distributions; unless specified otherwise, a distribution is represented by its cumulative distribution function (CDF).

Definition 1 (Index). Box i 's index θ_i is the unique solution to

$$\mathbb{E}_{v_i \sim F_i}[(v_i - \theta_i)_+] = c_i \quad \text{where} \quad (x)_+ \triangleq \max\{x, 0\} . \quad (\text{Index Eq.})$$

Definition 2 (The Index Algorithm, [Weitzman, 1978](#)). Compute indices for all boxes. Discard boxes with negative indices; sort the rest in decreasing order by their indices. Inspect the boxes in this order until the largest observed value exceeds the indices of all the boxes not inspected, or when all the boxes have been inspected. Take the box with the highest revealed value.

Both our observations ((a) and (b) above) are consequences of the discovery, by [Kleinberg et al. \(2016\)](#), that the searcher implementing the index algorithm always ends up taking the box with the highest *amortized value*:

Definition 3 (Amortized value). For box i with index θ_i and realized value v_i , its amortized value is $\kappa_i \triangleq \min\{v_i, \theta_i\}$.

Proposition 1 ([Kleinberg et al., 2016](#)). A searcher implementing the index algorithm in the Pandora Box problem takes the box with the highest nonnegative amortized value.

By Proposition 1, for a seller in a competitive search-good market, the competition from rival sellers can be summarized by one number: the highest amortized value from these sellers. This number is abstracted as o , the value of the outside option, in our setup in Section 2.1. Our seller in question is chosen by the buyer if and only if his own amortized value is greater than o .

When there is signaling, the buyer calculates the index using the posterior distribution, and chooses the sender iff his posterior amortized value is greater than o .

2.3 The Sender's Persuasion Problem

Given the receiver's optimal search rule, i.e., the Index Algorithm, the sender can reshape the receiver's belief about the invisible item and affect the corresponding index computed by the receiver through information disclosure. Now, we formalize the sender's persuasion problem.

A signaling scheme is a joint distribution over realized values and signals. Under the Index Algorithm, both the sender's and the receiver's utilities are determined by the realized values and the posterior indices. Since signals inducing the same posterior index are payoff-equivalent, we identify information strategies with joint distributions G over the value-index pair (v, θ) .

Definition 4 (Feasible joint distribution). A joint distribution $G \in \Delta([0, 1] \times [-c, 1 - c])$ is feasible if its value marginal is F and, for G_θ -almost every $\theta \in [-c, 1 - c]$,³

$$\int_0^1 ((v - \theta)_+ - c) \, dG(v \mid \theta) = 0 .$$

We denote the set of feasible joint distributions by $\mathcal{G}(F, c)$, which is convex and weakly compact.⁴

Our original game is payoff-equivalent to the following reduced game:

1. The sender chooses a feasible joint distribution $G \in \mathcal{G}(F, c)$ as his strategy.
2. The sender's value-index pair (v, θ) is drawn from G ; its amortized value κ is $\min\{v, \theta\}$.
3. The outside option o is drawn from H . If $\kappa \geq o$, the receiver takes the sender's item; otherwise, she takes the outside option.

³ For such a joint distribution G , we write G_θ for its marginal distribution over indices. We write $G_{\theta|v}$ and $G_{v|\theta}$ for the conditional distributions of θ given v and of v given θ , respectively.

⁴ See Appendix A.1 for the formal statement and the proof.

For the sender's utility, he gains utility 1 if chosen by the receiver, and utility 0 otherwise. For a realized amortized value $\kappa \in [-c, 1 - c]$, the sender's interim utility is the probability of the realized outside option o being smaller than κ , so we define

$$u(\kappa) \triangleq \begin{cases} 0 & \text{if } \kappa \in [-c, 0) , \\ H(\kappa) & \text{if } \kappa \in [0, 1 - c] . \end{cases}$$

The sender's persuasion problem is to choose an information strategy G among the joint distribution space $\mathcal{G}(F, c)$ to maximize the expected utility:

$$\max_{G \in \mathcal{G}(F, c)} U(G) \triangleq \mathbb{E}_{(v, \theta) \sim G} [u(\min\{v, \theta\})] . \quad (\mathcal{P}_{\text{BR}})$$

It is an infinite-dimensional linear program (LP) over joint value-index distributions, subject to a continuum of feasibility constraints. Consequently, neither standard finite-dimensional LP methods nor conventional one-dimensional persuasion techniques apply directly.

3 Optimal Strategy Verification

In this section, we develop a dual-certificate approach to the sender's problem. In Section 3.1, we show the existence of optimal solutions for both the primal and dual problems (Lemma 1), and derive a version of complementary slackness conditions (Lemma 2). In Section 3.2, we develop a method for deciding if a given primal solution is optimal (Theorem 1); this is by constructing a dual solution (Algorithms 1 and 3). Finally, we apply the framework to a competitive extension and certify a symmetric equilibrium for two senders with a common convex prior (Theorem 2).

3.1 Duality and Optimality Certificates

We formulate the sender's problem ($\mathcal{P}_{\text{BR}}$) as the primal problem. Define $p(v, \theta) \triangleq u(\min\{v, \theta\})$ and $q(v, \theta) \triangleq (v - \theta)_+ - c$ for each value-index pair (v, θ) . To derive the dual problem, we introduce a dual variable $\lambda(v)$ for the value-marginal constraint at each $v \in [0, 1]$, and a dual variable $\mu(\theta)$ for the index constraint at each $\theta \in [-c, 1 - c]$. Then, the dual problem is

$$\begin{aligned} \min_{\lambda, \mu} \quad & \int_0^1 \lambda(v) f(v) \, dv \\ \text{s.t.} \quad & \lambda(v) + \mu(\theta) q(v, \theta) \geq p(v, \theta) , \quad \forall (v, \theta) \end{aligned} \quad (\mathcal{D}_{\text{BR}})$$

For continuous interim utility function $u(\kappa)$, the primal must have an optimal solution G^* by the weak compactness of the feasible space $\mathcal{G}(F, c)$ (Lemma 4). We next show that, under mild assumptions, the dual also has an optimal solution, with additional well-behaved properties.

Lemma 1 (Existence of optimal dual solutions). *If function $u(\kappa)$ is L -Lipschitz continuous for some $L > 0$,⁵ the dual problem ($\mathcal{D}_{\text{BR}}$) has an optimal solution (λ^*, μ^*) such that (i) $\mu^*(\theta) \in [-L, 0]$ for each $\theta \in [-c, 1 - c]$; and (ii) λ^* is non-decreasing and continuous over $[0, 1]$.*

Such properties make the subsequent construction of dual variables tractable. Complementary slackness also holds, and can be used to verify whether a given pair of feasible solutions is optimal both to the primal and dual problems.

Lemma 2 (Complementary slackness). *Given a feasible primal solution G^* and a feasible dual solution (λ^*, μ^*) , G^* and (λ^*, μ^*) are both optimal to the primal and dual if and only if*

$$\int_{[0, 1] \times [-c, 1 - c]} [\lambda^*(v) + \mu^*(\theta) q(v, \theta) - p(v, \theta)] \, dG^*(v, \theta) = 0 . \quad (2)$$

⁵ Let $\mathcal{X}$ be a subset of a Euclidean space, a function $h : \mathcal{X} \rightarrow \mathbb{R}$ is L -Lipschitz continuous for some $L > 0$ if $|h(x) - h(y)| \leq L\|x - y\|$ for all $x, y \in \mathcal{X}$.

The lemma links primal supports with binding dual constraints. Dual feasibility implies that the slack $\lambda^*(v) + \mu^*(\theta)q(v, \theta) - p(v, \theta)$ is nonnegative everywhere. Hence, Condition (2) implies that this slack must vanish G^* -almost everywhere. Equivalently, for G^* -almost every pair (v, θ) ,

$$\lambda^*(v) = \sup_{\theta' \in [-c, 1-c]} \{p(v, \theta') - \mu^*(\theta')q(v, \theta')\} = p(v, \theta) - \mu^*(\theta)q(v, \theta), \quad (3)$$

where $p(v, \theta)$ is the payoff of the value-index pair (v, θ) , and $-\mu^*(\theta)q(v, \theta)$ captures its spillover effect on the other values with the same index. Thus, $p(v, \theta) - \mu^*(\theta)q(v, \theta)$ can be viewed as the amortized payoff generated by (v, θ) , while $\lambda^*(v)$ is the maximal amortized payoff among all indices. The optimal strategy assigns mass only to indices that maximize this amortized payoff.

3.2 Optimality Verification Procedures

Using techniques above, we develop a method to verify the optimality of a given primal solution. Suppose the prior F is atomless and strictly increasing over $[0, 1]$.

We first derive a necessary condition for optimality to rule out some candidates. Recall that $u(\kappa) = 0$ whenever $\kappa < 0$. Thus low values within $[0, c]$ can generate positive utility only if they are pooled with higher values to induce nonnegative amortized values. However, values sufficiently close to zero are not worth pooling: raising them to nonnegative amortized values requires too much distortion of higher values relative to the payoff they can generate. The following lemma formalizes this observation through a threshold value.

Lemma 3 (Positive threshold). *For any feasible strategy $G \in \mathcal{G}(F, c)$, define*

$$T \triangleq \left\{ v \in [0, c] \mid \begin{array}{l} \text{supp}(G_{v|\theta}) \cap [0, v) = \emptyset \text{ for all } \theta \in (v - c, 1 - c], \\ \text{supp}(G_{v|\theta}) \cap (v, 1] = \emptyset \text{ for all } \theta \in [-c, v - c] \end{array} \right\}.$$

The threshold $\underline{v} \triangleq \sup T$ if $T \neq \emptyset$; and $\underline{v} \triangleq 0$ otherwise. If G^ is optimal, then $\underline{v} > 0$.*

For candidate strategies that survive, we propose Algorithms 1 and 3, which construct dual variables λ and μ respectively (Algorithm 3 is in Appendix B.4). If the two algorithms terminate successfully, the input and their are both optimal for the respective programs.

Theorem 1 (Optimality verification). *Assume that: (i) $u(\kappa)$ is differentiable and Lipschitz continuous everywhere; and (ii) for every $(v, \theta) \in \text{supp}(G)$, any open neighborhood of (v, θ) has strictly positive probability under G . If $\underline{v} > 0$ and a feasible solution (λ, μ) can be constructed by Algorithms 1 and 3, then G and (λ, μ) are optimal for the primal and dual programs, respectively.*

Specifically, Algorithm 1 determines $\lambda(v)$ for each $v \in [0, 1]$ using spatial relationships between the dual variables, and Algorithm 3 checks the existence of $\mu(\theta)$ that constitutes a feasible dual solution along with λ constructed. Algorithm 3 is relatively straightforward, so we focus on explaining the construction of λ in the main text. A few definitions are in order.

Definition 5 (Monotone support sequence). *Given a joint distribution G , a monotone support sequence of G is a sequence $\{(v^m, \theta^m)\}_{m \in \mathbb{Z}_+} \subseteq \text{supp}(G)$ such that $v^m < v^{m+1}$ for every $m \in \mathbb{Z}_+$, and the sequence $\{\theta^m\}_{m \in \mathbb{Z}_+}$ is either nondecreasing or nonincreasing. If $\{\theta^m\}_{m \in \mathbb{Z}_+}$ is strictly monotone, then $\{(v^m, \theta^m)\}_{m \in \mathbb{Z}_+}$ is called a strictly monotone support sequence.*

The construction of λ is based on necessary conditions for G and (λ, μ) to be optimal. In other words, the deductions in the following are driven by the assumption that G is optimal, and that a corresponding λ satisfying the complementary slackness condition with G can be constructed. We provide an illustration for Algorithm 1 in the right panel of Figure 1. By our previous discussion following the complementary slackness conditions, if G and (λ, μ) are both optimal, $\lambda(v)$ represents value v 's contribution to the utility. For any $v \in [0, 1]$, the determination of $\lambda(v)$ falls into one of five cases. In the first two cases, we can directly determine $\lambda(v)$:

ALGORITHM 1: LAMBDA (v): Construction of the dual variable λ .**Input** : Candidate strategy G ; interim utility u ; cost c ; threshold $\underline{v}$; and any value $v \in [0, 1]$.**Output**: The corresponding value of the dual variable $\lambda(v)$.

```

if  $v \in [0, \underline{v}]$  then
  |  $\lambda(v) \leftarrow 0$  and return  $\lambda(v)$  ; // Case 1
end
if  $(v, v - c) \in \text{supp}(G)$  then
  |  $\lambda(v) \leftarrow u(v - c)$  and return  $\lambda(v)$  ; // Case 2
end
if there is a monotone support sequence converging to  $(v, v - c)$  then
  |  $\lambda(v) \leftarrow u(v - c)$  and return  $\lambda(v)$  ; // Case 3.1
end
else if there is a strictly monotone support sequence converging to  $(v, \theta_1)$  with  $\theta_1 \neq v - c$  then
  |  $\lambda(v) \leftarrow u'(\theta_1)q(v, \theta_1) + p(v, \theta_1)$  and return  $\lambda(v)$  ; // Case 3.2
end
else
  | There is a  $\theta_2 \neq v - c$  such that  $(v - \epsilon, v) \subseteq \text{supp}(G_{v|\theta_2})$  for some  $\epsilon > 0$  ;
  |  $\underline{v}_{\theta_2} \leftarrow \inf \text{supp}(G_{v|\theta_2})$  ;
  |  $\lambda(v) \leftarrow \frac{\text{LAMBDA}(\underline{v}_{\theta_2}) - p(\underline{v}_{\theta_2}, \theta_2)}{q(\underline{v}_{\theta_2}, \theta_2)} q(v, \theta_2) + p(v, \theta_2)$  and return  $\lambda(v)$  ; // Case 3.3
end

```

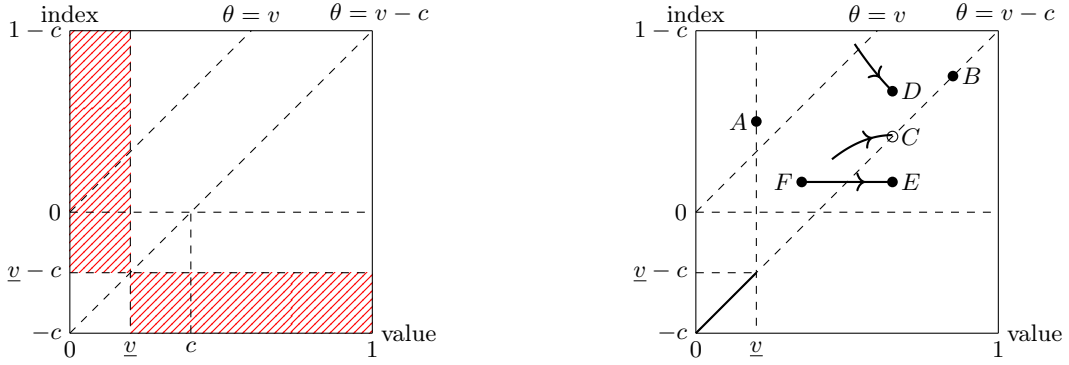


Fig. 1. The left panel shows the threshold value $\underline{v}$. The right panel illustrates Algorithm 1. The solid dots and black curves represent the support set of one candidate strategy G . The hollow dots represent the limit points of some monotone support sequence. Case 1: Let $\lambda(v_A) = 0$ since $v_A \leq \underline{v}$; Case 2: Let $\lambda(v_B) = u(v_B - c)$ since $\theta_B = v_B - c$; Case 3.1: Let $\lambda(v_C) = u(v_C - c)$ since there is a monotone support sequence converging to (v_C, θ_C) ; Case 3.2: Let $\lambda(v_D) = u'(\theta_D)q(v_D, \theta_D) + p(v_D, \theta_D)$ since there is a strictly monotone support sequence converging to (v_D, θ_D) ; Case 3.3: Let $\lambda(v_E) = \frac{\lambda(v_F) - p(v_F, \theta_F)}{q(v_F, \theta_F)} q(v_E, \theta_E) + p(v_E, \theta_E)$ since $(v_F, \theta_F) \subseteq \text{supp}(G_{v|\theta_E})$.

- **Case 1:** If $v \in [0, \underline{v}]$, then Algorithm 1 returns $\lambda(v) = 0$. By Lemma 3, values at or below the threshold $\underline{v}$ make no contribution to the sender's payoff.
- **Case 2:** If $(v, v - c) \in \text{supp}(G)$, then Algorithm 1 returns $\lambda(v) = u(v - c)$. Condition $\lambda(v) + \mu(v - c)q(v, v - c) - p(v, v - c) = 0$ implies $\lambda(v) = u(v - c)$ since $q(v, v - c) = 0$.

Now consider a value v that falls into neither of these two cases. As the prior distribution is assumed to have positive density everywhere on $[0, 1]$, there must exist a monotone support sequence converging to the point (v, θ) for some $\theta \in [0, 1 - c]$. Depending on the convergence properties of this sequence, we further divide the analysis into three cases:

- **Case 3.1:** If there is a monotone support sequence $\{(v^m, \theta^m)\}_{m \in \mathbb{Z}^+}$ converging to the point $(v, v - c)$, by the continuity of λ and the boundedness of μ (see Lemma 1), we have $\lambda(v) = \lim_{m \rightarrow \infty} -\mu(\theta^m)q(v^m, \theta^m) + p(v^m, \theta^m) = u(v - c)$.
- **Case 3.2:** If there is a strictly monotone support sequence $\{(v^m, \theta^m)\}_{m \in \mathbb{Z}^+}$ converging to the point (v, θ_1) for some $\theta_1 \neq v - c$. By Lemma 2, $\lambda(v^m) = -\mu(\theta^m)q(v^m, \theta^m) + p(v^m, \theta^m)$ for

- each $m \in \mathbb{Z}^+$. The converging sequence gives infinitely many such equalities in the neighborhood of (v, θ_1) . These equalities, combined with the specific formula of $-\mu(\theta^m)q(v^m, \theta^m) + p(v^m, \theta^m)$, allows us to conclude that $\lambda(v) = u'(\theta)q(v, \theta_1) + p(v, \theta_1)$.
- **Case 3.3:** If none of the above conditions hold, there must exist $\theta_2 \neq v - c$ such that $(v - \epsilon, v) \subseteq \text{supp}(G_{v|\theta_2})$ for some $\epsilon > 0$. By the continuity of λ , we reduce the task of solving $\lambda(v)$ to solving $\lambda(\inf \text{supp}(G_{\cdot|\theta_2}))$ and run Algorithm 1 again to solve this case.

3.3 Application: Certifying Symmetric Equilibrium

We now use the verification framework to certify symmetric equilibria in a competitive version of the model. We first extend the single-sender model to a multi-sender game, which can be viewed as a Pandora Box problem where each box is controlled by a strategic sender.

Competitive information design game. There are $n \geq 2$ senders, each controlling one item. All senders share the same prior F and cost $c > 0$. Each sender i chooses an information strategy $G_i \in \mathcal{G}(F, c)$. The value-index pair (v_i, θ_i) is independently drawn from G_i with the amortized value $\kappa_i \triangleq \min\{v_i, \theta_i\}$. The receiver selects a sender with the highest nonnegative amortized value. We denote the amortized value distribution under some strategy G by

$$K_G(\kappa) \triangleq \Pr[\min\{v, \theta\} \leq \kappa] = G(\kappa, 1 - c) + G(1, \kappa) - G(\kappa, \kappa) .$$

If all opponents use G , the induced interim utility of an individual sender is⁶

$$u_G(\kappa) \triangleq \begin{cases} 0 & \text{if } \kappa < 0 , \\ K_G(\kappa)^{N-1} & \text{if } \kappa \geq 0 . \end{cases}$$

A feasible strategy $G^* \in \mathcal{G}(F, c)$ forms a *symmetric equilibrium* if and only if G^* is a best response strategy to the competitive environment u_{G^*} .

Notice that [Lyu \(2023\)](#) also studies competition but only in the large-market limit. His dimensionality-reduction approach can not characterize finite-seller markets, which is precisely where our verification framework working on the joint distributions over (v, θ) becomes useful.

We consider a game where all senders share a common convex prior F .⁷ There are two cases depending on the inspection cost. When $c \leq \inf \text{supp}(F)$, all realized values induce nonnegative full-revelation indices, and full revelation is the unique symmetric equilibrium; this case has a simpler characterization, so we defer it to [Appendix B.6](#). We focus here on the high-cost case, where $c > \inf \text{supp}(F)$. For a two-sender game with a common convex prior F , we construct a symmetric equilibrium whose induced amortized-value distribution has a hinge-convex structure.

Theorem 2 (Hinge-convex signaling equilibrium). *For two senders with a common convex prior F and cost $c > \inf \text{supp}(F)$, there is a symmetric equilibrium G^* that induces each sender's amortized value distribution K of the following hinge-convex structure: for certain parameters $\theta_1 \in (-c, 0)$, $\theta_2 \in (0, 1 - c]$, and $\rho > 0$ (uniquely determined in the proof),*

$$K^*(\kappa) = \begin{cases} F(\kappa + c) & \text{if } \kappa \in [-c, \theta_1] , \\ F(\theta_1 + c) & \text{if } \kappa \in (\theta_1, 0] , \\ \min\{\rho \cdot \kappa + F(\theta_1 + c), 1\} & \text{if } \kappa \in (0, \theta_2] , \\ F(\kappa + c) & \text{if } \kappa \in (\theta_2, 1 - c] , \end{cases}$$

where $K^*(\theta_2) = F(\theta_2 + c)$ if $\theta_2 < 1 - c$ (the non-degenerate case); and $K^*(\theta_2) = 1$ if $\theta_2 = 1 - c$ (the degenerate case).

⁶ This expression ignores ties since ties cannot exist under equilibrium in most scenarios we consider.

⁷ We also provide results about the symmetric equilibrium under a common concave prior; see [Appendix E](#).

In this equilibrium, values below $\theta_1 + c$ are fully revealed and yield zero utility. Values in $(\theta_1 + c, c)$ are pooled with higher values to generate positive amortized values, which are spread uniformly over an interval starting from 0. In the non-degenerate case, pooling stops at value $\theta_2 + c$, and values above are fully revealed (see Figure 2). In the degenerate case, all values above $\theta_1 + c$ join the pooling region, so the upper full-revelation region disappears.

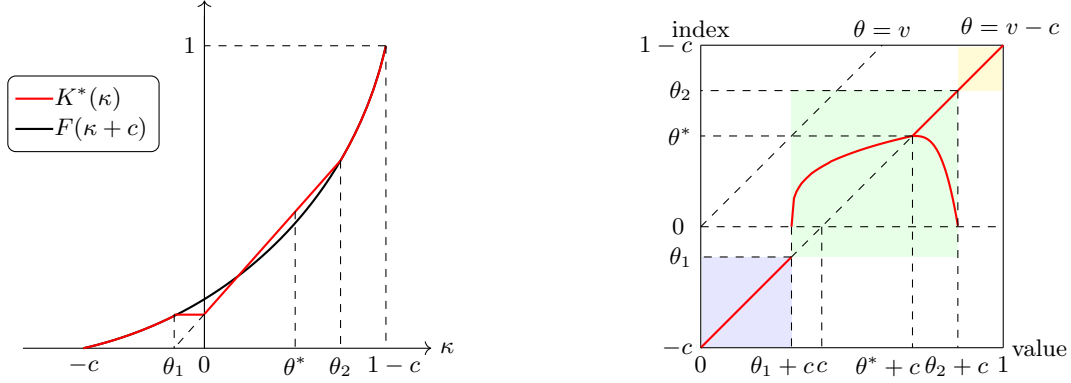


Fig. 2. An example of the non-degenerate case in Theorem 2. In the left panel, the red curve represents the amortized value distribution $K^*(\kappa)$ in equilibrium, and the black curve the shifted prior $F(\kappa + c)$. In the right panel, the red curves represent the support sets of the equilibrium strategy G^* . In the blue and yellow regions, low values and high values are fully revealed; while in the green region, lower values in $(\theta_1 + c, \theta^* + c)$ are paired with higher values in $(\theta^* + c, \theta_2 + c)$ to produce amortized values uniformly distributed over an interval.

Our findings echo a common pattern in real-world advertising. Firms often use such a *three-fold advertising regime*: they highlight the product (by truthfully revealing) when the consumer–product match is clearly good, stay silent when the match is clearly poor, and otherwise blur the match quality—benefiting most when rivals face similarly mediocre matches.

Our proof follows the *identify–construct–verify* procedure. We first identify the candidate equilibrium amortized-value distribution K^* , then construct a feasible joint distribution G^* that implements it, and finally construct dual variables (λ^*, μ^*) via Algorithms 1 and 3 and invoke Theorem 1 to certify that the constructed strategy is a best response to itself.

Step 1: Identify the candidate amortized value distribution K^* . For the target structure of the amortized value distribution K^* in Theorem 2, values in $(\theta_1 + c, \theta_2 + c]$ are pooled so that the induced amortized values are uniformly distributed over $[0, \theta'_1]$ for some parameter $\theta'_1 \leq \theta_2$ to be determined. These thresholds are pinned down by the following two equalities.

First, notice that the amortized value distribution K of any feasible joint distribution forms a mean-preserving contraction (MPC) of the shifted prior $F(\kappa + c)$.⁸ Due to this observation, it holds $K^*(\theta_1) = F(\theta_1 + c)$ and $K^*(\theta'_1) = F(\theta'_1 + c)$. Because K^* is assessed to be linear with slope ρ on this region, we obtain the first equality:

$$\rho\theta'_1 = F(\theta'_1 + c) - F(\theta_1 + c) .$$

Second, consider the boundary value $\theta_1 + c$. At equilibrium, this value must be indifferent between being pooled with higher values to produce amortized value 0 and being truthfully revealed to produce amortized value $\theta_1 < 0$. Thus, the gain from pooling, $F(\theta_1 + c)$, must equal the loss from lowering higher values' amortized values by $|\theta_1|$. Since K^* is linear with slope ρ in the pooled region, this loss is $\rho|\theta_1|$, giving the second equality:

$$F(\theta_1 + c) = \rho|\theta_1| .$$

⁸ A distribution $H \in \Delta([0, 1])$ forms an MPC of a distribution $F \in \Delta([0, 1])$ if and only if $\int_0^t F(x) dx \geq \int_0^t H(x) dx$ for any $t \in [0, 1]$ with equality at $t = 1$. See Lemma 3.1 in Lyu (2023) for a formal statement.

We finally prove that there uniquely exist θ_1 and θ'_1 that solve the resulting system obtained by combining the above two equalities.

$$\frac{F(\theta_1 + c)}{|\theta_1|} = \frac{F(\theta'_1 + c) - F(\theta_1 + c)}{\theta'_1} = \rho.$$

The solution yields two cases: If $K^*(\theta'_1) < 1$, the equilibrium is non-degenerate with $\theta_2 = \theta'_1$; otherwise, it is degenerate with $\theta_2 = 1 - c$. We sketch the remaining proof for the more involved non-degenerate case.

Step 2: Construct a joint distribution G^* that implements K^* . To implement the target amortized-value distribution K^* , we fully reveal values below $\theta_1 + c$ and above $\theta_2 + c$, and split the intermediate region $[\theta_1 + c, \theta_2 + c]$ into low and high parts. Since the density f is nondecreasing and $F(\theta_2 + c) - F(\theta_1 + c) = K^*(\theta_2) - K^*(\theta_1) = \rho\theta_2$, there uniquely exists a cutoff $\theta^* \in (\theta_1, \theta_2)$ such that each value above $\theta^* + c$ has density of size at least ρ . We call such values *high*, and the remaining values in the interval *low*.

High values serve two roles in the construction. Each high value v is truthfully revealed with probability $\rho/f(v)$, and its remaining mass is pooled with low values. The pairing must simultaneously generate the desired amortized value distribution and preserve both the value-marginal constraint and the row-wise index constraints.

We construct the pairing through a system of ODEs. Each index $\theta \in [0, \theta^*]$ pools one low value $\alpha(\theta)$ with one high value $\beta(\theta)$. The initial condition pairs the lowest and highest values: $\alpha(0) = \theta_1 + c$ and $\beta(0) = \theta_2 + c$. The ODEs are chosen so that each row satisfies the index equation and the induced amortized values have the target density. Their solution satisfies $\alpha'(\theta) \geq 0$ and $\beta'(\theta) \leq 0$ for all $\theta \in [0, \theta^*]$, and terminates at $\alpha(\theta^*) = \beta(\theta^*) = \theta^* + c$.

Step 3: Certify equilibrium via the verification framework. It remains to show that the constructed strategy is indeed a best response to itself. This is the point at which the verification framework becomes essential. Rather than directly solving the sender's best-response problem, given the interim utility induced by the hinge-shaped amortized-value distribution K^* , we construct dual functions (λ^*, μ^*) using Algorithms 1 and 3. We then invoke Theorem 1 to certify the optimality of the constructed strategy and dual functions.

4 Optimal Strategy Approximation

In this section, we analyze the sender's best-response problem from a complementary perspective. We provide an additive FPTAS to compute a near-optimal strategy (Theorem 3).

Theorem 3 (An additive FPTAS). *Suppose the interim utility function $u(\kappa)$ is L -Lipschitz continuous for some $L > 0$. For every $\epsilon > 0$, Algorithm 2 computes a feasible strategy $\tilde{G} \in \mathcal{G}(F, c)$ whose payoff is at least $\text{OPT} - \epsilon$ in polynomial time in L and $1/\epsilon$. In other words, the sender's persuasion problem ($\mathcal{P}_{\text{BR}}$) admits a fully polynomial-time approximation scheme (FPTAS).*

The theorem complements the verification framework developed above. Since we keep the joint distributions over value-index pairs as the decision variables, we have to preserve both the value marginal and the row-wise index constraints throughout the approximation.

Our approximation scheme uses a *two-level* discretization. We first discretize the value space to obtain a discrete prior, and then discretize the index support to obtain a finite LP (see Figure 3 for illustrations). The proof compares three optimization problems: the original continuous-prior problem, the discretized-prior problem, and the finite LP after index discretization. We solve for the finite LP and lift the discrete optimal solution back to the feasible region under the original prior as the output of our algorithm. The key step is to bound the payoff loss in each transition between the three problems above. We provide the proof sketch below.

Proof Sketch. The first problem is the original sender's persuasion problem ($\mathcal{P}_{\text{BR}}$), which admits an optimal solution G^* with optimal value OPT .

The second problem is obtained by discretizing the value prior. For a grid size $\delta \in (0, 1)$, we uniformly partition the value space $[0, 1]$ into intervals $\{I_r\}_{r \in [M]}$. We then move all probability mass of F in each interval I_r to its midpoint v_r , thereby obtaining a discrete prior

$$F_\delta \triangleq \sum_{i=1}^M f_i^\delta \cdot \delta_{v_i} \quad \text{where} \quad f_i^\delta \triangleq F(I_i), \quad (4)$$

and δ_{v_i} denotes the Dirac measure at v_i . We define the optimal value under the optimum G_δ^*

$$\text{OPT}_\delta \triangleq \max_{G_\delta \in \mathcal{G}(F_\delta, c)} U(G_\delta) = \mathbb{E}_{(v, \theta) \sim G_\delta} [u(\min\{v, \theta\})] . \quad (\mathcal{P}_{\text{BR}}^\delta)$$

The two problems are not directly comparable: their feasible sets are $\mathcal{G}(F, c)$ and $\mathcal{G}(F_\delta, c)$, neither of which contains the other. Thus, there is no immediate ordering between their optimal values. Nevertheless, their optimal values satisfy the following two-sided stability bound:

$$|\text{OPT} - \text{OPT}_\delta| \leq L\delta . \quad (5)$$

To prove this, we construct feasible solutions in both directions. Starting from G^* , we move each row's value mass to the corresponding grid midpoints and recompute the row index, obtaining a feasible solution under F_δ . Since both values and recomputed indices move by at most δ , the Lipschitz continuity of u implies a payoff loss of at most $L\delta$. Conversely, starting from G_δ^* , we spread each midpoint mass back to its original value interval according to the conditional prior and recompute row indices. The same argument also gives a payoff loss of at most $L\delta$.

The third problem is a finite LP obtained by further discretizing the index support. Let $v_1 < \dots < v_M$ denote the support of the discretized prior F_δ . Define $\Theta_\delta \triangleq [v_1 - c, v_M - c]$ as the effective index domain under F_δ . Fix a granularity $\epsilon \in (0, 1)$, we first discretize the payoff range of the interim utility $u(\kappa)$. For each $\ell_k \triangleq k\epsilon$, we compute the corresponding generalized inverse $\xi_k \triangleq \inf\{\theta \in \Theta_\delta : u(\theta) \geq \ell_k\}$. We then identify the full index grids:

$$\Gamma \triangleq \{v_i - c : i \in [M]\} \cup (\{v_i : i \in [M]\} \cap \Theta_\delta) \cup \{\xi_k : k = 0, 1, \dots, \lceil 1/\epsilon \rceil\} . \quad (6)$$

Denote $R \triangleq |\Gamma|$. The points v_i ensure that the mapping $\theta \mapsto \min\{v_i, \theta\}$ has no kink inside any grid interval. The points $v_i - c$ ensure that the function $q_i(\theta) = (v_i - \theta)_+ - c$ does not change sign inside any grid interval. The payoff-grid points ξ_k ensure that moving mass between adjacent grid indices causes at most ϵ utility loss. Let $\mathbf{x} = (x_{ir})_{i \in [M], r \in [R]}$ be the decision variable, where $x_{ir} \geq 0$ denotes the probability mass assigned to (v_i, θ_r) . We consider the following finite LP:

$$\begin{aligned} \text{OPT}_{\delta, \epsilon} &\triangleq \max_{\mathbf{x} \geq 0} \sum_{i=1}^M \sum_{r=1}^R u(\min\{v_i, \theta_r\}) x_{ir} \\ \text{s.t.} \quad &\sum_{r=1}^R x_{ir} = f_i^\delta, \quad \forall i \in [M], \\ &\sum_{i=1}^M ((v_i - \theta_r)_+ - c) x_{ir} = 0, \quad \forall r \in [R]. \end{aligned} \quad (\mathcal{P}_{\text{BR}}^{\delta, \epsilon})$$

Since every solution is also feasible for the discretized-prior problem ($\mathcal{P}_{\text{BR}}^\delta$), it holds $\text{OPT}_{\delta, \epsilon} \leq \text{OPT}_\delta$. The converse approximation follows from a rounding argument. Any feasible strategy under F_δ can be rounded to the grid Γ by splitting each non-grid row between the two adjacent grid rows. The construction ensures that the total loss is at most ϵ . Hence

$$\text{OPT}_\delta - \epsilon \leq \text{OPT}_{\delta, \epsilon} \leq \text{OPT}_\delta . \quad (7)$$

Finally, we solve the finite LP and lift its optimal solution $G_{\delta,\epsilon}^*$ back to the original prior F . By the same lifting argument as above, this step incurs at most $L\delta$ payoff loss. Combining the loss bounds (5) and (7) achieved above, the resulting feasible strategy $\tilde{G} \in \mathcal{G}(F, c)$ satisfies

$$U(\tilde{G}) \geq \text{OPT}_{\delta,\epsilon} - L\delta \geq \text{OPT} - \epsilon - 2L\delta.$$

Thus, the constructed strategy $\tilde{G}$ is the output of our FPTAS, which achieves any constant factor approximation to the optimal solution G^* .

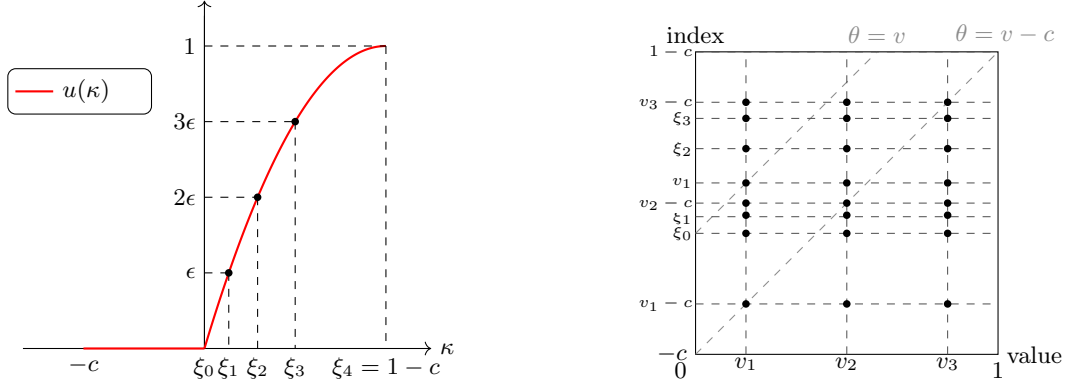


Fig. 3. Illustration of the two-level discretization in Theorem 3. The left panel shows the payoff-grid construction for the interim utility $u(\kappa)$. The right panel shows the resulting finite linear program supported on the value-index grids after discretizations. The black points denote the supports of the discrete solution under the finite LP.

ALGORITHM 2: An additive FPTAS

Input : Prior $F \in \Delta([0, 1])$; Inspection cost $c > 0$; L -Lipschitz interim utility $u(\kappa)$; Target error $\bar{\epsilon} \in (0, 1)$.

Output: A feasible strategy $\tilde{G} \in \mathcal{G}(F, c)$ with payoff at least $\text{OPT} - \bar{\epsilon}$.

Set the discretization parameters $\epsilon \leftarrow \bar{\epsilon}/2$ and $\delta \leftarrow \min\{1/2, \bar{\epsilon}/4L\}$

Construct the discretized prior F_δ according to (4) under parameter δ

Construct the index grids I according to (6) under parameter ϵ

Construct and solve the finite linear program $(\mathcal{P}_{\text{BR}}^{\delta,\epsilon})$. Let $G_{\delta,\epsilon}^*$ be an optimal solution

Lift $G_{\delta,\epsilon}^*$ back to the original prior F by spreading each mass at v_i over I_i , and obtain $\tilde{G}$

return $\tilde{G}$

Bibliography

- Simon P Anderson and Régis Renault. Advertising content. *American Economic Review*, 96(1):93–113, 2006.
- Simon P Anderson and Régis Renault. The advertising mix for a search good. *Management Science*, 59(1):69–83, 2013.
- Mark Armstrong. Ordered consumer search. *Journal of the European Economic Association*, 15(5):989–1024, 2017.
- Pak Hung Au and Mark Whitmeyer. Attraction versus persuasion: Information provision in search markets. *Journal of Political Economy*, 131(1):202–245, 2023.
- Santiago Balseiro, Yuan Deng, Jieming Mao, Vahab Mirrokni, and Song Zuo. Optimal mechanisms for a value maximizer: The futility of screening targets. In *Proceedings of the 25th ACM Conference on Economics and Computation*, EC ’24, page 1101, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400707049. <https://doi.org/10.1145/3670865.3673574>. URL <https://doi.org/10.1145/3670865.3673574>.
- Kiarash Banihashem, Mohammad Taghi Hajiaghayi, Piotr Krysta, and Suho Shin. Delegated choice with combinatorial constraints. In *Proceedings of the 26th ACM Conference on Economics and Computation*, pages 275–303, 2025.
- Heski Bar-Isaac, Guillermo Caruana, and Vicente Cuñat. Search, design, and market structure. *American Economic Review*, 102(2):1140–1160, 2012.
- Dirk Bergemann and Stephen Morris. Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95, 2019.
- Hedyeh Beyhaghi and Linda Cai. Pandora’s problem with nonobligatory inspection: Optimal structure and a PTAS. In Barna Saha and Rocco A. Servedio, editors, *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 803–816. ACM, 2023a.
- Hedyeh Beyhaghi and Linda Cai. Pandora’s problem with nonobligatory inspection: Optimal structure and a PTAS. In Barna Saha and Rocco A. Servedio, editors, *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 803–816. ACM, 2023b.
- Hedyeh Beyhaghi and Robert Kleinberg. Pandora’s problem with nonobligatory inspection. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 131–132, 2019.
- Simon Board and Jay Lu. Competitive information disclosure in search markets. *Journal of Political Economy*, 126(5):1965–2010, 2018.
- Raphael Boleslavsky, Christopher S Cotton, and Haresh Gurnani. Demonstrations and price competition in new product release. *Management Science*, 63(6):2016–2026, 2017.
- Shant Boodaghians, Federico Fusco, Philip Lazos, and Stefano Leonardi. Pandora’s box problem with order constraints. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 439–458, 2020.
- Shuchi Chawla, Evangelia Gergatsouli, Yifeng Teng, Christos Tzamos, and Ruimin Zhang. Pandora’s box with correlations: Learning and approximation. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1214–1225. IEEE, 2020.
- Michael Choi, Kyungmin Kim, and Marilyn Pease. Optimal information design for search goods. In *AEA Papers and Proceedings*, volume 109, pages 550–556. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, 2019.
- Yuan Deng, Mohammad Mahdian, Jieming Mao, Vahab Mirrokni, Hanrui Zhang, and Song Zuo. Efficiency of the generalized second-price auction for value maximizers. In *Proceedings of the ACM Web Conference 2024, WWW ’24*, page 46–56, New York, NY, USA, 2024.

- Association for Computing Machinery. ISBN 9798400701719. <https://doi.org/10.1145/3589334.3645360>. URL <https://doi.org/10.1145/3589334.3645360>.
- Mahsa Derakhshan, Negin Golrezaei, Vahideh H. Manshadi, and Vahab Mirrokni. Product ranking on online platforms. *Manag. Sci.*, 68(6):4024–4041, 2022.
- Bolin Ding, Yiding Feng, Chien-Ju Ho, Wei Tang, and Haifeng Xu. Competitive information design for pandora’s box. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 353–381. SIAM, 2023.
- Mustafa Dogan and Ju Hu. Consumer search and optimal information. *The RAND Journal of Economics*, 53(2):386–403, 2022.
- Laura Doval. Whether or not to open pandora’s box. *Journal of Economic Theory*, 175:127–158, 2018.
- Shaddin Dughmi. Algorithmic information structure design: a survey. *ACM SIGecom Exchanges*, 15(2):2–24, 2017.
- Piotr Dworczak and Anton Kolotilin. The persuasion duality. *Theoretical Economics*, 19(4):1701–1755, 2024.
- Piotr Dworczak and Giorgio Martini. The simple economics of optimal persuasion. *Journal of Political Economy*, 127(5):1993–2048, 2019.
- Ophir Friedler, Hu Fu, Anna R. Karlin, and Ariana Tang. Price stability and improved buyer utility with presentation design: A theoretical study of the amazon buy box. In Guodong Long, Michale Blumstein, Yi Chang, Liane Lewin-Eytan, Zi Helen Huang, and Elad Yom-Tov, editors, *Proceedings of the ACM on Web Conference 2025, WWW 2025, Sydney, NSW, Australia, 28 April 2025- 2 May 2025*, pages 4592–4611. ACM, 2025.
- Hu Fu, Jiawei Li, and Daogao Liu. Pandora box problem with nonobligatory inspection: Hardness and approximation scheme. In Barna Saha and Rocco A. Servedio, editors, *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 789–802. ACM, 2023.
- Mohammad Taghi Hajiaghayi, Piotr Krysta, Mohammad Mahdavi, and Suho Shin. Delegation with costly inspection. In *Proceedings of the 26th ACM Conference on Economics and Computation, EC ’25*, page 871–894, New York, NY, USA, 2025. Association for Computing Machinery. ISBN 9798400719431. <https://doi.org/10.1145/3736252.3742640>. URL <https://doi.org/10.1145/3736252.3742640>.
- Toomas Hinnosaar and Keiichi Kawai. Robust pricing with refunds. *The RAND Journal of Economics*, 51(4):1014–1036, 2020.
- Dongjin Hwang and Ilwoo Hwang. Competitive information disclosure with heterogeneous consumer search. *arXiv preprint arXiv:2504.04659*, 2025.
- Ilwoo Hwang, Kyungmin Kim, and Raphael Boleslavsky. Competitive advertising and pricing. *Review of Economic Studies*, page rdag062, 2026.
- Maxim Ivanov. Information revelation in competitive markets. *Economic Theory*, 52(1):337–365, 2013.
- Emir Kamenica. Bayesian persuasion and information design. *Annual Review of Economics*, 11(1):249–272, 2019.
- Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.
- Robert Kleinberg, Bo Waggoner, and E Glen Weyl. Descending price optimally coordinates search. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pages 23–24, 2016.
- Anton Kolotilin. Optimal information disclosure: A linear programming approach. *Theoretical Economics*, 13(2):607–635, 2018.
- Chen Lyu. Information design for selling search goods and the effect of competition. *Journal of Economic Theory*, 213:105722, 2023.
- Martin Weitzman. *Optimal search for the best alternative*, volume 78. Department of Energy, 1978.

- Mark Whitmeyer. Persuasion produces the (diamond) paradox. *arXiv preprint arXiv:2011.13900*, 2020.
- Christopher A. Wilkens, Ruggiero Cavallo, and Rad Niazadeh. Mechanism design for value maximizers. *CoRR*, abs/1607.04362, 2016. URL <http://arxiv.org/abs/1607.04362>.
- Christopher A. Wilkens, Ruggiero Cavallo, and Rad Niazadeh. GSP - the cinderella of mechanism design. *CoRR*, abs/1701.05946, 2017. URL <http://arxiv.org/abs/1701.05946>.

A Proofs in Section 2

A.1 Lemma 4 and Its Proof

Lemma 4. *For any prior F and cost c , the action space $\mathcal{G}(F, c)$ is convex and weakly compact.*

Proof of Lemma 4. We divide the proof into two steps.

First, we show that the strategy space $\mathcal{G}(F, c)$ is convex given any prior F and cost c . Consider two feasible strategies $G \in \mathcal{G}(F, c)$ and $H \in \mathcal{G}(F, c)$ (with densities g and h). We want to prove that the convex combination $K(x) = \lambda \cdot G(x) + (1 - \lambda) \cdot H(x)$ for any $x \in [0, 1]$ (with density $\lambda \cdot g + (1 - \lambda) \cdot h$) also belongs to space $\mathcal{G}(F, c)$ for any $\lambda \in [0, 1]$.

To see this, first, for all indices $\theta \in [-c, 1 - c]$, it holds

$$\lambda \cdot \mathbb{E}_{v \sim G_{|\theta}}[\max\{v - \theta, 0\}] + (1 - \lambda) \cdot \mathbb{E}_{v \sim H_{|\theta}}[\max\{v - \theta, 0\}] = c .$$

Besides, for each value $v \in [0, 1]$, it holds

$$\lambda \cdot \int_{\theta=-c}^{1-c} g_{|v}(\theta) \, d\theta + (1 - \lambda) \cdot \int_{\theta=-c}^{1-c} h_{|v}(\theta) \, d\theta = f(v) .$$

Therefore, we have proved that any convex combination of the two feasible strategies also satisfies all the constraints of space $\mathcal{G}(F, c)$, which makes $\mathcal{G}(F, c)$ convex.

Second, we show that the strategy space $\mathcal{G}(F, c)$ is also compact. The boundedness of space $\mathcal{G}(F, c)$ can be directly obtained by the fact that space $\mathcal{G}(F, c)$ is indeed a measure space over $[0, 1] \times [-c, 1 - c]$. Since $[0, 1] \times [-c, 1 - c]$ is a compact metric space, the space of probability measures $\Delta([0, 1] \times [-c, 1 - c])$ is compact in the weak topology by Prokhorov's theorem.

Next, we prove that the strategy space $\mathcal{G}(F, c)$ is weak closed. We assume that there exists a sequence $\{G^m\}_{m \in \mathbb{Z}}$ with each $G^m \in \mathcal{G}(F, c)$, weakly converges to a certain 2-D distribution G , we shall prove the limit point G also belongs to set $\mathcal{G}(F, c)$.

To see this, we first check the marginal constraint (Bayesian plausibility). Since each $G^m \in \mathcal{G}(F, c)$, we have

$$\int_{\theta=-c}^{1-c} g^m(v, \theta) \, d\theta = f(v) , \quad \forall v \in [0, 1] .$$

It is equivalent to the following condition: for all bounded and continuous function ϕ defined on $[0, 1]$, we have

$$\int_{(v, \theta)} \phi(v) \, dG^m(v, \theta) = \int_{v=0}^1 \phi(v) \, dF(v) .$$

According to the definition of weak convergence, we have

$$\lim_{m \rightarrow \infty} \int_{(v, \theta)} \phi(v) \, dG^m(v, \theta) = \int_{(v, \theta)} \phi(v) \, dG(v, \theta) = \int_{v=0}^1 \phi(v) \, dF(v) ,$$

which is further equivalent to

$$\int_{\theta=-c}^{1-c} g(v, \theta) \, d\theta = f(v) , \quad \forall v \in [0, 1] .$$

We then check the index constraint, which is

$$\int_{v=0}^1 ((v - \theta)_+ - c) g^m(v, \theta) \, dv = 0 , \quad \forall \theta \in [-c, 1 - c] .$$

It is equivalent to the following condition: for all bounded and continuous function ψ defined on $[-c, 1 - c]$, we have

$$\begin{aligned} & \int_{\theta=-c}^{1-c} \psi(\theta) \left(\int_{v=0}^1 ((v - \theta)_+ - c) dG^m(v, \theta) \right) d\theta \\ &= \int_{(v, \theta)} \psi(\theta) ((v - \theta)_+ - c) dG^m(v, \theta) = 0. \end{aligned}$$

Both functions ψ and $(v - \theta)_+ - c$ are bounded and continuous in θ and (v, θ) , respectively. According to the definition of weak convergence, we have

$$\lim_{m \rightarrow \infty} \int_{(v, \theta)} \psi(\theta) ((v - \theta)_+ - c) dG^m(v, \theta) = \int_{(v, \theta)} \psi(\theta) ((v - \theta)_+ - c) dG(v, \theta) = 0.$$

This fact is equivalent to

$$\int_{v=0}^1 ((v - \theta)_+ - c) g(v, \theta) dv = 0, \quad \forall \theta \in [-c, 1 - c].$$

Thus we have proved that the limit point G must belong to space $\mathcal{G}(F, c)$, which makes the space closed. To sum up, we have achieved the boundedness and closeness of space $\mathcal{G}(F, c)$ under weak convergence, and we can directly make space $\mathcal{G}(F, c)$ weak compact through the Heine-Borel Theorem. $\square$

B Proofs in Section 3

B.1 Proof of Lemma 1

Before proving Lemma 1, we have to introduce the following Lemmas 5 to 7 with respect to the properties of dual variables $\lambda(v)$ and $\mu(\theta)$.

Lemma 5. *For any dual solution (λ, μ) , it holds that $\lambda(v) \geq u(v - c) \geq 0$ for each $v \in [0, 1]$.*

Proof of Lemma 5. For each $v \in [0, 1]$, by the fact that tuple $(v, v - c)$ satisfies Condition (1), we directly have $\lambda(v) \geq u(v - c) \geq 0$. $\square$

Lemma 6. *Assume there exists a constant $L > 0$ such that the interim utility function u is L -Lipschitz continuous over $[-c, 1 - c]$. If there exists an optimal solution (λ^*, μ^*) to the dual problem $\mathcal{D}_{BR}$, we can without loss assume that $\mu^*(\theta) \in [-L, 0]$ for each $\theta \in [-c, 1 - c]$.*

Proof of Lemma 6. First, we prove $\mu^*(\theta) \leq 0$ for each $\theta \in [-c, 1 - c]$ by contradiction. We assume that there exists an optimal dual solution (λ^*, μ^*) with $\mu^*(\bar{\theta}) > 0$ for some $\bar{\theta} \in [-c, 1 - c]$. If $\lambda^*(v) > p(v, \bar{\theta}) - \mu^*(\bar{\theta})q(v, \bar{\theta})$ holds for all $v \in [\bar{\theta}, 1]$, then assuming $\mu^*(\bar{\theta}) = 0$ is without loss of generality, since this assumption does not affect the dual objective while keeping the constraints hold. If there exists a value $\bar{v} > \bar{\theta}$ with $\lambda^*(\bar{v}) = p(\bar{v}, \bar{\theta}) - \mu^*(\bar{\theta})q(\bar{v}, \bar{\theta})$, then for any $\theta \geq \bar{v}$, by Constraint (1), it holds that $\lambda^*(\bar{v}) \geq u(\bar{v})$. For the tuple $(\bar{v}, \bar{\theta})$, also by Constraint (1) and the monotonicity of u , we know that $\lambda^*(\bar{v}) = u(\bar{\theta}) - \mu^*(\bar{\theta})(\bar{v} - \bar{\theta}) < u(\bar{\theta}) < u(\bar{v})$, which forms a contradiction. To sum up, we have proved that, for any optimal solution (λ^*, μ^*) to the dual problem, we can without loss assume that $\mu^*(\theta) \leq 0$ for each $\theta \in [-c, 1 - c]$.

Second, we prove that for each $\theta \in [-c, 1 - c]$, $\mu^*(\theta)$ is bounded by the constant L . The Lipschitz continuity of function u implies that $|u(x) - u(y)| \leq L \cdot |x - y|$ for any $x, y \in [-c, 1 - c]$. We assume that there exists a $\theta_0 \in (-c, 1 - c)$ (the case of $\theta_0 = -c$ or $1 - c$ is trivial) with $\mu(\theta_0) < -L$. Then we know that $u(v - c) \geq -\mu(\theta_0)q(v, \theta_0) + p(v, \theta_0)$ for any $v \in [0, \theta_0 + c]$ and $u(v - c) \leq -\mu(\theta_0)q(v, \theta_0) + p(v, \theta_0)$ for any $v \in [\theta_0 + c, 1]$. So setting $\mu(\theta) = -L$ will weakly further decrease the objective function. To sum up, we have proved that $-\mu(\theta) \leq L$, that is function μ is bounded by the constant L . Directly by Constraint (1), we know that the function λ corresponding to the bounded function μ , is also bounded. $\square$

Lemma 7. *Assume there exists a constant $L > 0$ such that the interim utility function u is L -Lipschitz continuous over $[-c, 1 - c]$. If there exists an optimal solution (λ^*, μ^*) to the dual problem $\mathcal{D}_{\text{BR}}$, then the function $\lambda^*(\cdot)$ is non-decreasing and continuous over $[0, 1]$.*

Proof of Lemma 7. Suppose (λ^*, μ^*) is optimal to the dual problem $\mathcal{D}_{\text{BR}}$. By Constraint (1) and the objective of the dual problem, we have

$$\lambda^*(v) = \max_{\theta \in [-c, 1-c]} -\mu^*(\theta)q(v, \theta) + p(v, \theta), \quad \forall v \in [0, 1].$$

Given a certain $\theta \in [-c, 1 - c]$, for all $v \in [0, 1]$, we define

$$\hat{L}_\theta(v) \triangleq -\mu^*(\theta)q(v, \theta) + p(v, \theta) = \begin{cases} -\mu^*(\theta)(v - \theta - c) + u(\theta) & \text{if } v \geq \theta, \\ u(v) + c\mu^*(\theta) & \text{if } v < \theta. \end{cases}$$

We observe that function $\hat{L}_\theta(\cdot)$ is linear over $[\theta, 1]$ and forms a shifted version of function u over $[0, \theta]$. Besides, $\lambda^*(v) = \max_{\theta \in [-c, 1-c]} \hat{L}_\theta(v)$ for each $v \in [0, 1]$. By Lemma 6, it is obvious that function $\hat{L}_\theta$ is non-decreasing over $[0, 1]$ for each $\theta \in [-c, 1 - c]$. By the fact that $\lambda^*(v) = \max_{\theta \in [-c, 1-c]} \hat{L}_\theta(v)$ for each $v \in [0, 1]$, we have the function $\lambda^*(\cdot)$ is also non-decreasing over $[0, 1]$. Since the dual variable μ^* is a bounded function (Lemma 6) and the function u is continuous over $[-c, 1 - c]$, we know that each $\hat{L}_\theta(\cdot)$ is continuous over $[0, 1]$ for any $\theta \in [-c, 1 - c]$. This further implies that the function $\lambda^*(\cdot)$ is also continuous over $[0, 1]$. $\square$

With these lemmas above, now we can prove Lemma 1.

Proof of Lemma 1. Based on the necessary conditions of the optimal solution to the dual problem (Lemmas 6 and 7), the dual problem $\mathcal{D}_{\text{BR}}$ is equivalent to the following convex optimization problem $\mathcal{D}_{\text{NEW}}$:

$$\begin{aligned} \min_{\lambda, \mu} \quad & \int_0^1 \lambda(v)f(v) dv & (\mathcal{D}_{\text{NEW}}) \\ \text{subject to} \quad & \lambda(v) = \max_{\theta \in [-c, 1-c]} \{p(v, \theta) - \mu(\theta)q(v, \theta)\}, \quad \forall v \in [0, 1] \end{aligned} \quad (8)$$

$$\mu(\theta) \leq 0, \quad \forall \theta \in [-c, 1 - c] \quad (9)$$

Any optimal solution (λ^*, μ^*) to the problem $\mathcal{D}_{\text{NEW}}$ also forms an optimal solution to the dual problem $\mathcal{D}_{\text{BR}}$, and vice versa. Thus, it suffices to prove that $\mathcal{D}_{\text{NEW}}$ has an optimal solution, and for any optimal solution (λ^*, μ^*) it holds: (i) $\mu^*(\theta) \in [-L, 0]$ for any $\theta \in [-c, 1 - c]$; and (ii) λ^* is non-decreasing and continuous over $[-c, 1 - c]$. We divide the whole proof into two steps.

Step-1: There exists an optimal solution (λ^*, μ^*) to the problem $\mathcal{D}_{\text{NEW}}$. Let

$$\text{OPT} = \inf_{\lambda, \mu} \int_0^1 \lambda(v)f(v) dv$$

denote the optimal value of the dual problem. Let $\text{OBJ}(\lambda, \mu)$ denote the objective value of any pair of feasible solutions (λ, μ) . We aim to show that there exists a feasible solution (λ^*, μ^*) with $\text{OBJ}(\lambda^*, \mu^*) = \text{OPT}$. Whether the dual problem has an optimal solution or not, there always exists a feasible solution sequence $\{(\lambda^m, \mu^m)\}_{m \in \mathbb{Z}^+}$ such that the corresponding objective sequence $\{\text{OBJ}(\lambda^m, \mu^m)\}_{m \in \mathbb{Z}^+}$ converges to the optimal value OPT^* .

By Lemma 7, we know that function λ is non-decreasing. Besides, without loss of generality, we can assume that function λ is also bounded. Combining these two facts, the sequence $\{\lambda^m\}_{m \in \mathbb{Z}^+}$ also weakly converges to certain function λ^* that satisfies $\int_0^1 \lambda^*(v)f(v) dv = \text{OPT}^*$. Next, we show that there exists a function μ^* that forms a feasible solution with the function

λ^* . For any $m \in \mathbb{Z}^+$, each pair (λ^m, μ^m) is feasible. So by Constraint (1), for any $\theta \in [-c, 1-c]$, we have that

$$\max_{v \in [\theta+c, 1]} \frac{u(\theta) - \lambda^m(v)}{v - \theta - c} \leq \mu^m(\theta) \leq \min \left\{ \min_{v \in [0, \theta]} \frac{\lambda^m(v) - u(v)}{c}, \min_{v \in [\theta, \theta+c]} \frac{u(\theta) - \lambda^m(v)}{v - \theta - c} \right\}.$$

Since the pair (λ^m, μ^m) is feasible, we have

$$\max_{v \in [\theta+c, 1]} \frac{u(\theta) - \lambda^m(v)}{v - \theta - c} \leq \min \left\{ \min_{v \in [0, \theta]} \frac{\lambda^m(v) - u(v)}{c}, \min_{v \in [\theta, \theta+c]} \frac{u(\theta) - \lambda^m(v)}{v - \theta - c} \right\},$$

and such μ^m exists. When $m \rightarrow \infty$, we have

$$\lim_{m \rightarrow \infty} \max_{v \in [\theta+c, 1]} \frac{u(\theta) - \lambda^m(v)}{v - \theta - c} = \max_{v \in [\theta+c, 1]} \frac{u(\theta) - \lambda^*(v)}{v - \theta - c},$$

and we have

$$\begin{aligned} & \lim_{m \rightarrow \infty} \min \left\{ \min_{v \in [0, \theta]} \frac{\lambda^m(v) - u(v)}{c}, \min_{v \in [\theta, \theta+c]} \frac{u(\theta) - \lambda^m(v)}{v - \theta - c} \right\} \\ &= \min \left\{ \min_{v \in [0, \theta]} \frac{\lambda^*(v) - u(v)}{c}, \min_{v \in [\theta, \theta+c]} \frac{u(\theta) - \lambda^*(v)}{v - \theta - c} \right\}. \end{aligned}$$

Thus, we know that there also exists such μ^* that forms a feasible solution along with function λ^* . To sum up, we have proved that there exists an optimal solution (λ^*, μ^*) to the problem $\mathcal{D}_{\text{NEW}}$. By the equivalence of $\mathcal{D}_{\text{NEW}}$ and $\mathcal{D}_{\text{BR}}$, there also exists an optimal solution to the dual problem $\mathcal{D}_{\text{BR}}$.

Step-2: For any optimal solution (λ^*, μ^*) to the dual problem $\mathcal{D}_{\text{BR}}$, $\mu^*(\theta) \in [-L, 0]$ for any $\theta \in [-c, 1-c]$, and λ^* is non-decreasing and continuous over $[-c, 1-c]$. These conditions hold directly by Lemmas 6 and 7. $\square$

B.2 Proof of Lemma 2

Proof of Lemma 2. First, we prove the sufficiency of Lemma 2. The dual variables $\lambda(v)$ and $\mu(\theta)$ are both bounded measurable functions, so the order of integration can be interchanged for any double integral over $[0, 1] \times [-c, 1-c]$ under Fubini's Theorem. It is obvious that both the primal and the dual have a feasible solution. We consider any pair of feasible solutions (g, λ, μ) . First, by the value-marginal constraint, we have that

$$\int_{v=0}^1 \int_{\theta=-c}^{1-c} \lambda(v) g(v, \theta) dv d\theta = \int_0^1 \lambda(v) f(v) dv. \quad (10)$$

By the row-wise index constraint, we have that

$$\int_{v=0}^1 \int_{\theta=-c}^{1-c} \mu(\theta) q(v, \theta) g(v, \theta) dv d\theta = \int_{\theta=-c}^{1-c} \int_{v=0}^1 \mu(\theta) q(v, \theta) g(v, \theta) dv d\theta = 0. \quad (11)$$

After combining Equation (10) and Equation (11), we achieve that

$$\int_{v=0}^1 \int_{\theta=-c}^{1-c} [\lambda(v) + \mu(\theta) q(v, \theta)] g(v, \theta) dv d\theta = \int_0^1 \lambda(v) f(v) dv. \quad (12)$$

By Constraint (1), we have that

$$\int_{v=0}^1 \int_{\theta=-c}^{1-c} [\lambda(v) + \mu(\theta) q(v, \theta)] g(v, \theta) dv d\theta \geq \int_{v=0}^1 \int_{\theta=-c}^{1-c} p(v, \theta) g(v, \theta) dv d\theta. \quad (13)$$

Next, we assume that there exists a pair of feasible solutions (g^*, λ^*, μ^*) that satisfies Condition (2), that is

$$\int_{v=0}^1 \int_{\theta=-c}^{1-c} [\lambda^*(v) + \mu^*(\theta)q(v, \theta)] g^*(v, \theta) \, dv \, d\theta = \int_{v=0}^1 \int_{\theta=-c}^{1-c} p(v, \theta) g^*(v, \theta) \, dv \, d\theta . \quad (14)$$

For this special pair of feasible solutions, after combining Equation (12) and Equation (14), then we have

$$\int_0^1 \lambda^*(v) f(v) \, dv = \int_{v=0}^1 \int_{\theta=-c}^{1-c} p(v, \theta) g^*(v, \theta) \, dv \, d\theta . \quad (15)$$

For any feasible solution g to the primal problem, after combining Equation (12) and Equation (13), we obtain that

$$\int_0^1 \lambda^*(v) f(v) \, dv \geq \int_{v=0}^1 \int_{\theta=-c}^{1-c} p(v, \theta) g(v, \theta) \, dv \, d\theta . \quad (16)$$

By Equation (15) and Inequality (16), we have

$$\int_{v=0}^1 \int_{\theta=-c}^{1-c} p(v, \theta) g^*(v, \theta) \, dv \, d\theta \geq \int_{v=0}^1 \int_{\theta=-c}^{1-c} p(v, \theta) g(v, \theta) \, dv \, d\theta ,$$

which means the special solution g^* is optimal to the primal problem ($\mathcal{P}_{\text{BR}}$). In the same manner, we can also prove that (λ^*, μ^*) is the optimal solution to the dual problem ($\mathcal{D}_{\text{BR}}$). Besides, Equation (15) implies that the strong duality holds between this primal problem and the dual problem.

Then we prove the necessity of Lemma 2. We assume that G^* with density g^* is an optimal solution to ($\mathcal{P}_{\text{BR}}$) and (λ^*, μ^*) is the optimal solution to $\mathcal{D}_{\text{BR}}$. Since the strong duality holds, we know that

$$\int_0^1 \lambda^*(v) f(v) \, dv = \int_{v=0}^1 \int_{\theta=-c}^{1-c} p(v, \theta) g^*(v, \theta) \, dv \, d\theta . \quad (17)$$

By the value-marginal constraint, we have that

$$\int_{v=0}^1 \int_{\theta=-c}^{1-c} \lambda^*(v) g^*(v, \theta) \, dv \, d\theta = \int_0^1 \lambda^*(v) f(v) \, dv . \quad (18)$$

By the row-wise index constraint, we have

$$\int_{v=0}^1 \int_{\theta=-c}^{1-c} \mu^*(\theta) q(v, \theta) g^*(v, \theta) \, dv \, d\theta = 0 . \quad (19)$$

Combining Equation (18) and Equation (19), we have achieve that

$$\int_{v=0}^1 \int_{\theta=-c}^{1-c} [\lambda^*(v) + \mu^*(\theta)q(v, \theta)] g^*(v, \theta) \, dv \, d\theta = \int_0^1 \lambda^*(v) f(v) \, dv . \quad (20)$$

Combining Equation (17) and Equation (12), we achieve that

$$\int_{v=0}^1 \int_{\theta=-c}^{1-c} [\lambda^*(v) + \mu^*(\theta)q(v, \theta) - p(v, \theta)] g^*(v, \theta) \, dv \, d\theta = 0 ,$$

which means that the optimal solution to the primal problem G^* and the optimal solution to the dual problem (λ^*, μ^*) satisfy Condition (2). $\square$

B.3 Proof of Lemma 3

Proof of Lemma 3. For any feasible 2-D distribution G , there being no such threshold value $\underline{v}$ implies that, there exists a tuple $(0, \theta) \in \text{supp}(G)$ such that $\theta > 0$. In any best response G^* , without loss of generality, we can assume that such $(0, \theta) \in \text{supp}(G)$ with $\theta > 0$ does not exist since pooling value 0 up to a positive index brings no utility increase. $\square$

B.4 Algorithm 3 for Construction of Dual Variable $\mu(\theta)$

Here is the algorithm to construct the dual variable $\mu(\theta)$ for any $\theta \in [-c, 1 - c]$.

ALGORITHM 3: MU (θ): Constructions of the dual variable μ

Require: Strategy G ; Interim utility u ; Cost c ; Threshold value $\underline{v}$; All $\lambda(v)$ for $v \in [0, 1]$.

Ensure : Whether the construction is valid (True/False).

```

for  $\theta \in [-c, \underline{v} - c]$  do
  |  $\mu(\theta) \leftarrow 0$  ;
end
for  $\theta \in \text{supp}(G_\theta) \cup (\underline{v} - c, 1 - c]$  do
  | if  $\text{supp}(G_{\cdot|\theta}) = \{\theta + c\}$  then
  |   | if  $\max_{v \in (\theta + c, 1]} \frac{u(\theta) - \lambda(v)}{v - \theta - c} \leq \min \left\{ \min_{v \in [0, \theta]} \frac{\lambda(v) - u(v)}{c}, \min_{v \in [\theta, \theta + c)} \frac{u(\theta) - \lambda(v)}{v - \theta - c} \right\}$  then
  |   |   | return False ;
  |   | end
  |   |  $\mu(\theta) \leftarrow \min \left\{ \min_{v \in [0, \theta]} \frac{\lambda(v) - u(v)}{c}, \min_{v \in [\theta, \theta + c)} \frac{u(\theta) - \lambda(v)}{v - \theta - c} \right\}$  ;
  | end
  | else
  |   |  $\mu_{\text{ref}} \leftarrow \text{None}$  ;
  |   | for  $v \in \text{supp}(G_{\cdot|\theta}) \setminus \{\theta + c\}$  do
  |   |   |  $\mu_{\text{cur}} \leftarrow \frac{p(v, \theta) - \lambda(v)}{q(v, \theta)}$  ;
  |   |   | if  $\mu_{\text{ref}} = \text{None}$  then
  |   |   |   |  $\mu_{\text{ref}} \leftarrow \mu_{\text{cur}}$  ;
  |   |   | end
  |   |   | else if  $\mu_{\text{cur}} \neq \mu_{\text{ref}}$  then
  |   |   |   | return False ;
  |   |   | end
  |   | end
  |   |  $\mu(\theta) \leftarrow \mu_{\text{ref}}$  ;
  | end
end
for  $\theta \notin \text{supp}(G_\theta)$  do
  | if  $\max_{v \in (\theta + c, 1]} \frac{u(\theta) - \lambda(v)}{v - \theta - c} \leq \min \left\{ \min_{v \in [0, \theta]} \frac{\lambda(v) - u(v)}{c}, \min_{v \in [\theta, \theta + c)} \frac{u(\theta) - \lambda(v)}{v - \theta - c} \right\}$  then
  |   | return False ;
  | end
  |  $\mu(\theta) \leftarrow \min \left\{ \min_{v \in [0, \theta]} \frac{\lambda(v) - u(v)}{c}, \min_{v \in [\theta, \theta + c)} \frac{u(\theta) - \lambda(v)}{v - \theta - c} \right\}$  ;
end
return True ;

```

B.5 Proof of Theorem 1

Before proving Theorem 1, we have to introduce the following lemmas.

Lemma 8 (Case 1 in Algorithm 1). *Given G^* and (λ^*, μ^*) are both optimal to the primal and dual problems, then we without loss of generality assume that $\lambda^*(v) = 0$ for each $v \in [0, \underline{v}]$ and $\mu^*(\theta) = 0$ for each $\theta \in [-c, \underline{v} - c]$.*

Proof of Lemma 8. Since strategy G^* is optimal to the primal problem, the threshold value $\underline{v}$ is above zero. Assuming no pooling is made to values $[0, \underline{v})$ is without loss of generality, thus we have $\lambda^*(v) = 0$ for any $v \in [0, \underline{v}]$. Besides, we can assume that $\mu^*(\theta) = 0$ for each $\theta \in [-c, \underline{v} - c]$ without loss of generality. Since assuming $\mu^*(\theta) = 0$ for each $\theta \in [-c, \underline{v} - c]$ does not affect the objective of the dual while keeping all constraints held. $\square$

Lemma 9 (Case 2 in Algorithm 1). *Given G^* and (λ^*, μ^*) are both optimal to the primal and dual problems, if there exists $v \in [\underline{v}, 1]$ such that $(v, v - c) \in \text{supp}(G)$, then $\lambda^*(v) = u(v - c)$.*

Proof of Lemma 9. By Lemma 2, it directly holds that $\lambda^*(v) = -\mu^*(v - c)q(v, v - c) + p(v, v - c) = p(v, v - c) = u(v - c)$ since $q(v, v - c) = 0$. $\square$

Lemma 10 (Case 3.1 in Algorithm 1). *Given G^* and (λ^*, μ^*) are both optimal to the primal and dual problems, for any (v, θ) with $v = \theta + c$, if there is a sequence $\{(v^m, \theta^m)\}_{m \in \mathbb{Z}^+}$ converging to (v, θ) such that : (i) $(v^m, \theta^m) \in \text{supp}(G)$ for any $m \in \mathbb{Z}^+$; (ii) $\{v^m\}_{m \in \mathbb{Z}^+}$ is strictly increasing; (iii) $\{\theta^m\}_{m \in \mathbb{Z}^+}$ is monotone; and (iv) $v^m \in (\theta^m, v)$ for any $m \in \mathbb{Z}^+$, then $\lambda^*(v) = u(v - c)$.*

Proof of Lemma 10. For any $m \in \mathbb{Z}^+$, since $(v^m, \theta^m) \in \text{supp}(G^*)$, Lemma 2 implies that

$$\lambda^*(v^m) = -\mu^*(\theta^m)q(v^m, \theta^m) + p(v^m, \theta^m) = -\mu^*(\theta^m)(v^m - \theta^m - c) + u(\theta^m).$$

Besides, since the function u is Lipschitz continuous everywhere, we have already proved the boundedness of the function u in Lemma 1. Thus, we know that

$$\lambda^*(v) = \lim_{m \rightarrow \infty} \lambda^*(v^m) = \lim_{m \rightarrow \infty} p(v^m, \theta^m) = u(v - c).$$

This completes the proof. $\square$

Lemma 11 (Case 3.2 in Algorithm 1 for the limit point (v, θ) with $v < \theta$). *Given G^* and (λ^*, μ^*) are both optimal to the primal and dual problems, for any (v, θ) with $v \in (\underline{v}, \theta)$, if there is a sequence $\{(v^m, \theta^m)\}_{m \in \mathbb{Z}^+}$ converging to (v, θ) such that : (i) $(v^m, \theta^m) \in \text{supp}(G)$ for any $m \in \mathbb{Z}^+$; (ii) $\{v^m\}_{m \in \mathbb{Z}^+}$ is strictly increasing; (iii) $\{\theta^m\}_{m \in \mathbb{Z}^+}$ is monotone; and (iv) $\theta^m > v$ for any $m \in \mathbb{Z}^+$, then $-\mu^*(\theta^{m_1}) = u'(\theta^{m_1}) = u'(\theta^{m_2}) = -\mu^*(\theta^{m_2})$ for any $m_1, m_2 \in \mathbb{Z}^+$. Besides, $\lambda^*(v) = u'(\theta^1)q(v, \theta) + p(v, \theta)$.*

Proof of Lemma 11. Without loss of generality, we can assume that $\{\theta^m\}_{m \in \mathbb{Z}^+}$ is also strictly increasing. For any $m \in \mathbb{Z}^+$, by Lemma 2 and the fact that function λ^* is continuous, we know that for any $m_1, m_2 \in \mathbb{Z}^+$ with $m_1 > m_2$,

$$\lambda^*(v^{m_1}) = (-c) \max_{\theta' \in [v, 1-c]} -\mu^*(\theta') + u(v^{m_1}) = -\mu^*(\theta^{m_1})(-c) + u(v^{m_1}),$$

and

$$\lambda^*(v^{m_2}) = (-c) \max_{\theta' \in [v, 1-c]} -\mu^*(\theta') + u(v^{m_2}) = -\mu^*(\theta^{m_2})(-c) + u(v^{m_2}).$$

Combining these two equations, we achieve that

$$\max_{\theta' \in [v, 1-c]} -\mu^*(\theta') = -\mu^*(\theta^{m_1}) = -\mu^*(\theta^{m_2}).$$

Next, we aim to prove that $\mu^*(\theta^{m_1}) = -\mu^*(\theta^{m_2}) = u'(\theta^{m_1}) = u'(\theta^{m_2})$. To form feasible indices θ^{m_1} and θ^{m_2} , there must exist $(\hat{v}^{m_1}, \theta^{m_1}) \in \text{supp}(G^*)$ with $\hat{v}^{m_1} > \theta^{m_1} + c$, and $(\hat{v}^{m_2}, \theta^{m_2}) \in \text{supp}(G^*)$ with $\hat{v}^{m_2} > \theta^{m_2} + c$. By Lemma 2, it holds that

$$\lambda^*(\hat{v}^{m_1}) = -\mu^*(\theta^{m_1})(\hat{v}^{m_1} - \theta^{m_1} - c) + u(\theta^{m_1}) \geq -\mu^*(\theta^{m_2})(\hat{v}^{m_1} - \theta^{m_2} - c) + u(\theta^{m_2}),$$

and

$$\lambda^*(\hat{v}^{m_2}) = -\mu^*(\theta^{m_2})(\hat{v}^{m_2} - \theta^{m_2} - c) + u(\theta^{m_2}) \geq -\mu^*(\theta^{m_1})(\hat{v}^{m_2} - \theta^{m_1} - c) + u(\theta^{m_1}).$$

Since (i) the function $-\mu^*(\theta^{m_1})q(\cdot, \theta^{m_1}) + u(\theta^{m_1})$ is linear over $[\theta_1, 1 - c]$ and crosses $(\theta^{m_1} + c, u(\theta^{m_1}))$; (ii) the function $-\mu^*(\theta^{m_2})q(\cdot, \theta^{m_2}) + u(\theta^{m_2})$ is linear over $[\theta_2, 1 - c]$ and crosses $(\theta^{m_2} + c, u(\theta^{m_2}))$; and (iii) $-\mu^*(\theta^{m_1}) = -\mu^*(\theta^{m_2})$, it must hold that $\mu^*(\theta^{m_1}) = -\mu^*(\theta^{m_2}) = u'(\theta^{m_1}) = u'(\theta^{m_2})$ and these two functions overlap over $[\theta_2 + c, 1 - c]$.

Recall that λ^*, p, q are all continuous; we complete our proof by showing that

$$\lambda^*(v) = \lim_{m \rightarrow \infty} \lambda^*(v^m) = \lim_{m \rightarrow \infty} -\mu^*(\theta^m)q(v^m, \theta^m) + p(v^m, \theta^m) = u'(\theta^1)q(v, \theta) + p(v, \theta).$$

This completes the proof. $\square$

Lemma 12 (Case 3.2 in Algorithm 1 for the limit point (v, θ) with $v \in (\theta, \theta + c)$).

Given G^* and (λ^*, μ^*) are both optimal to the primal and dual problems, for any (v, θ) with $v \in (\max\{\underline{v}, \theta\}, \theta + c)$, if there is a sequence $\{(v^m, \theta^m)\}_{m \in \mathbb{Z}^+}$ converging to (v, θ) such that: (i) $(v^m, \theta^m) \in \text{supp}(G)$ for any $m \in \mathbb{Z}^+$; (ii) $\{v^m\}_{m \in \mathbb{Z}^+}$ is strictly increasing; (iii) $\{\theta^m\}_{m \in \mathbb{Z}^+}$ is monotone; (iv) let $\theta' = \frac{v+\theta}{2}$, $\theta^m \in (v - c, \theta')$ and $v^m \in (\theta', v)$ for any $m \in \mathbb{Z}^+$, then $-\mu^*(\theta^{m_1}) = u'(\theta^{m_1}) = u'(\theta^{m_2}) = -\mu^*(\theta^{m_2})$ for any $m_1, m_2 \in \mathbb{Z}^+$. Besides, $\lambda^*(v) = u'(\theta^1)q(v, \theta) + p(v, \theta)$.

Proof of Lemma 12. Without loss of generality, we can assume that $\{\theta^m\}_{m \in \mathbb{Z}^+}$ is also strictly increasing. We can consider two tuples $(v^3, \theta^3), (v^4, \theta^4)$ with $\theta' \leq v_3 < v_4 < v$ and $\theta_4 > \theta_3$. There exist tuples $(\bar{v}^3, \theta^3), (\bar{v}^4, \theta^4) \in \text{supp}(G)$ such that $\bar{v}^3 > \theta^3 + c$ and $\bar{v}^4 > \theta^4 + c$. Based on Lemma 2, we achieve the following inequalities:

$$\begin{aligned} -\mu^*(\theta^3)q(v^3, \theta^3) + p(v^3, \theta^3) &\geq -\mu^*(\theta^4)q(v^3, \theta^4) + p(v^3, \theta^4), \\ -\mu^*(\theta^3)q(\bar{v}^3, \theta^3) + p(\bar{v}^3, \theta^3) &\geq -\mu^*(\theta^4)q(\bar{v}^3, \theta^4) + p(\bar{v}^3, \theta^4), \\ -\mu^*(\theta^4)q(v^4, \theta^4) + p(v^4, \theta^4) &\geq -\mu^*(\theta^3)q(v^4, \theta^3) + p(v^4, \theta^3), \\ -\mu^*(\theta^4)q(\bar{v}^4, \theta^4) + p(\bar{v}^4, \theta^4) &\geq -\mu^*(\theta^3)q(\bar{v}^4, \theta^3) + p(\bar{v}^4, \theta^3). \end{aligned}$$

Since (i) the function $-\mu^*(\theta^3)q(\cdot, \theta^3) + p(\cdot, \theta^3)$ is linear over (θ', v) ; (ii) the function $-\mu^*(\theta^4)q(\cdot, \theta^4) + p(\cdot, \theta^4)$ is linear over (θ', v) ; and (iv) the function $-\mu^*(\theta^4)q(\cdot, \theta^4) + p(\cdot, \theta^4)$ crosses $(\theta^4 + c, u(\theta^4))$, the above four inequalities imply that $\mu^*(\theta^3) = \mu^*(\theta^4) = u'(\theta^3) = u'(\theta^4)$ and these two functions overlap over $(\theta', 1]$. The above analysis holds for any $m_1, m_2 \in \mathbb{Z}^+$ with $m_1 < m_2$. Thus we know that $\mu^*(\theta^{m_1}) = \mu^*(\theta^{m_2}) = u'(\theta^{m_1}) = u'(\theta^{m_2})$ for any $m_1, m_2 \in \mathbb{Z}^+$. For any $m \in \mathbb{Z}^+$, it holds

$$\lambda^*(v^m) = p(v^m, \theta^m) - \mu^*(\theta^m)q(v^m, \theta^m).$$

Passing m to infinity, we achieve that

$$\lambda^*(v) = p(v, \theta) - \mu^*(\theta^1)q(v, \theta) = u'(\theta^1)q(v, \theta) + p(v, \theta).$$

This completes the proof. $\square$

Lemma 13 (Case 3.2 in Algorithm 1 for the limit point (v, θ) with $v > \theta + c$). Given G^* and (λ^*, μ^*) are both optimal to the primal and dual problems, for any (v, θ) with $v > \theta + c$, if there is a sequence $\{(v^m, \theta^m)\}_{m \in \mathbb{Z}^+}$ converging to (v, θ) such that: (i) $(v^m, \theta^m) \in \text{supp}(G)$ for any $m \in \mathbb{Z}^+$; (ii) $\{v^m\}_{m \in \mathbb{Z}^+}$ is strictly increasing; (iii) $\{\theta^m\}_{m \in \mathbb{Z}^+}$ is monotone; (iv) let $\theta' = \frac{v-c+\theta}{2} \in (\theta, v - c)$, $\theta^m \in [0, \theta')$ and $v^m \in (\theta' + c, v)$ for any $m \in \mathbb{Z}^+$, then $-\mu^*(\theta^{m_1}) = u'(\theta^{m_1}) = u'(\theta^{m_2}) = -\mu^*(\theta^{m_2})$ for any $m_1, m_2 \in \mathbb{Z}^+$. Besides, $\lambda^*(v) = u'(\theta^1)q(v, \theta) + p(v, \theta)$.

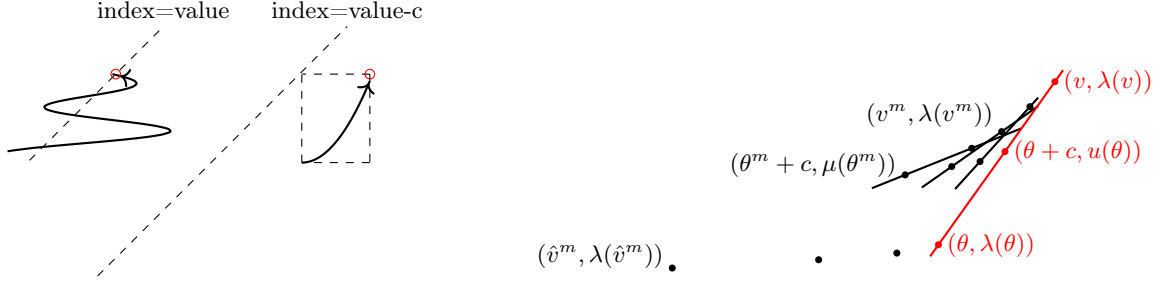


Fig. 4. Graph illustration of Lemma 13

Proof of Lemma 13. We proceed with the proof in two steps.

Step-1: Prove that $-\mu^*(\theta^{m_1}) = -\mu^*(\theta^{m_2}) = u'(\theta^{m_1}) = u'(\theta^{m_2})$ **for any** $m_1, m_2 \in \mathbb{Z}^+$. Let $v_l^m \triangleq \inf \text{supp}(G_{\cdot|\theta^m}^*)$ for any $m \in \mathbb{Z}^+$. Since the sequence $\{(v^m, \theta^m)\}_{m \in \mathbb{Z}^+}$ converges to (v, θ) , the sequence $\{(v_l^m, \theta^m)\}_{m \in \mathbb{Z}^+}$ also converges to some (v_l, θ) with $v_l < \theta + c$. We divide the proof into the following three cases.

- (i) **Case 1:** $v_l < \theta$. We can prove this result using a similar method as Lemma 11.
- (ii) **Case 2:** $v_l \in (\theta, \theta + c)$. We can prove this result using a similar method as Lemma 12.
- (iii) **Case 3:** $v_l = \theta$. If the sequence $\{\theta^m\}$ is also decreasing, then we know that there exists a subsequence $\{m(k)\}_{k \in \mathbb{Z}^+}$ such that $\{(v_l^{m(k)}, \theta^{m(k)})\}_{k \in \mathbb{Z}^+}$ converges to (v_l, θ) , and is decreasing. In this case, we can prove the target result using a similar method as Lemma 11.

Thus, without loss, we can assume that the sequence $\{\theta^m\}$ is increasing, and there exists a subsequence $\{m(k)\}_{k \in \mathbb{Z}^+}$ such that $\{(v_l^{m(k)}, \theta^{m(k)})\}_{k \in \mathbb{Z}^+}$ converges to (v_l, θ) , and is also increasing. Abusing the notations a little, we rename the sequence $\{m(k)\}_{k \in \mathbb{Z}^+}$ as $\{m\}_{m \in \mathbb{Z}^+}$. Since the sequences $\{v^m\}_{m \in \mathbb{Z}^+}$ and $\{\theta^m\}_{m \in \mathbb{Z}^+}$ are both increasing, by Lemma 2, we know that $\{-\mu^*(\theta^m)\}_{m \in \mathbb{Z}^+}$ is also increasing. We have already prove the boundedness of μ^* , thus $\{-\mu^*(\theta^m)\}_{m \in \mathbb{Z}^+}$ also converges to some real number, which is denoted as $-\hat{\mu}$, that is

$$\lim_{m \rightarrow \infty} -\mu^*(\theta^m) = -\hat{\mu}.$$

Next, we aim to prove that there exists an $M > 0$ such that for any $m > M$, $-\mu^*(\theta^m) = -\hat{\mu}$. We prove it through contradiction. We assume that for any $m < \infty$, it holds that $-\mu^*(\theta^m) < -\hat{\mu}$. We consider three sequences: (i) $\{(v_l^m, \lambda(v_l^m))\}_{m \in \mathbb{Z}^+}$, (ii) $\{(\theta^m + c, u(\theta^m))\}_{m \in \mathbb{Z}^+}$, and (iii) $\{(v^m, \lambda(v^m))\}_{m \in \mathbb{Z}^+}$. Let (x_a, y_a) denote the limit point of the sequence $\{(v_l^m, \lambda(v_l^m))\}_{m \in \mathbb{Z}^+}$. Let (x_b, y_b) denote the limit point of the sequence $\{(\theta^m + c, u(\theta^m))\}_{m \in \mathbb{Z}^+}$. Let (x_c, y_c) denote the limit point of the sequence $\{(v^m, \lambda(v^m))\}_{m \in \mathbb{Z}^+}$. Based on the graph illustrations, we can observe that $u'(\theta) \leq -\hat{\mu}$. Next we aim to show that these three points $(x_a, y_a), (x_b, y_b), (x_c, y_c)$ lie in a common linear function. First, we show that $\frac{y_c - y_b}{x_c - x_b} = -\hat{\mu}$:

$$\begin{aligned} \frac{y_c - y_b}{x_c - x_b} &= \lim_{m \rightarrow \infty} \frac{\lambda^*(v^m) - u(\theta^m)}{v^m - (\theta^m + c)} \\ &= \lim_{m \rightarrow \infty} \frac{-\mu^*(\theta^m)[v^m - (\theta^m + c)]}{v^m - (\theta^m + c)} \\ &= \lim_{m \rightarrow \infty} -\mu^*(\theta^m) = -\hat{\mu}. \end{aligned}$$

Next, we prove that $\frac{y_b - y_a}{x_b - x_a} = -\hat{\mu}$. For any $m \in \mathbb{Z}^+$, we have that

$$\begin{aligned} \frac{y_b - y_a}{x_b - x_a} &= \lim_{m \rightarrow \infty} \frac{u(\theta^m) - \lambda(v_l^m)}{\theta^m + c - v_l^m} \\ &= \lim_{m \rightarrow \infty} \frac{u(\theta^m) - (-\mu^*(\theta^m)q(v_l^m, \theta^m) + p(v_l^m, \theta^m))}{\theta^m + c - (v_l^m - \theta^m) - \theta^m} = -\hat{\mu}. \end{aligned}$$

Combining these two facts, we know that $(x_a, y_a), (x_b, y_b), (x_c, y_c)$ lie in a common linear function. Under this fact, we know that for sufficiently small $\epsilon > 0$,

$$-\mu^*(\theta - \epsilon)q(v_l, \theta - \epsilon) + p(v_l, \theta - \epsilon) > -\mu^*(\theta)q(v_l, \theta) + p(v_l, \theta) ,$$

which forms a contradiction with the fact that the sequence $\{v_l^m, \theta^m\}_{m \in \mathbb{Z}^+}$ converges to (v_l, θ) with $v_l = \theta$. Thus, we have proved that the assumption in the beginning is invalid. To sum up, we have proved that there exists a sufficiently large $M > 0$ such that $-\mu^*(\theta^{m_1}) = -\mu^*(\theta^{m_2})$ for any $m_1, m_2 > M$. Using a similar idea of Lemma 11, we can also prove that $-\mu^*(\theta^{m_1}) = -\mu^*(\theta^{m_2}) = u'(\theta^{m_1}) = u'(\theta^{m_2})$ for any $m_1, m_2 > M$.

Step-2: Prove that $\lambda^*(v) = u'(\theta^1)q(v, \theta) + p(v, \theta)$. For any $m \in \mathbb{Z}^+$, Lemma 2 implies that

$$\lambda^*(v^m) = p(v^m, \theta^m) - \mu^*(\theta^m)q(v^m, \theta^m) .$$

Passing m to infinity, since functions λ^*, p, q are both continuous, we achieve that

$$\lambda^*(v) = p(v, \theta) - \mu^*(\theta^1)q(v, \theta) = u'(\theta^1)q(v, \theta) + p(v, \theta) .$$

This completes the proof. $\square$

Lemma 14 (Case 3.2 in Algorithm 1 for the limit point (v, θ) with $v = \theta$). *Given G^* and (λ^*, μ^*) are both optimal to the primal and dual problems, for any (v, θ) with $v = \theta + c$, if there is a sequence $\{(v^m, \theta^m)\}_{m \in \mathbb{Z}^+}$ converging to (v, θ) such that : (i) $(v^m, \theta^m) \in \text{supp}(G)$ for any $m \in \mathbb{Z}^+$; (ii) $\{v^m\}_{m \in \mathbb{Z}^+}$ is strictly increasing; and (iii) $\{\theta^m\}_{m \in \mathbb{Z}^+}$ is monotone, then there exists a sufficiently large $M > 0$ such that $-\mu^*(\theta^m) = u'(\theta)$ for any $m > M$. Besides, $\lambda^*(v) = u'(\theta)q(v, \theta) + p(v, \theta)$.*

Proof of Lemma 14. Based on the properties of the sequence $\{(v^m, \theta^m)\}_{m \in \mathbb{Z}^+}$, we can divide the proof into three cases:

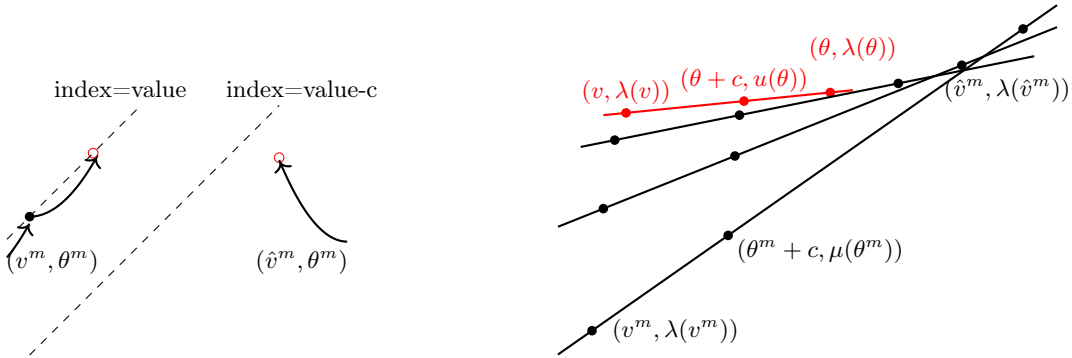


Fig. 5. Graph illustration of Lemma 14

Case 1: The sequence $\{\theta^m\}_{m \in \mathbb{Z}^+}$ is decreasing. This lemma can be proved in a method similar to that of Lemma 11.

Case 2: The sequence $\{\theta^m\}_{m \in \mathbb{Z}^+}$ is increasing, and there exists $M > 0$ such that $v^m \geq \theta^m$ for any $m > M$. For any $m \in \mathbb{Z}^+$, there exists $(\hat{v}^m, \theta^m) \in \text{supp}(G^*)$ with $\hat{v}^m > \theta^m + c$. By Lemma 2, we have that

$$p(\hat{v}^m, \theta^m) - \mu^*(\theta^m)q(\hat{v}^m, \theta^m) \geq p(\hat{v}^m, \theta^{m+1}) - \mu^*(\theta^{m+1})q(\hat{v}^m, \theta^{m+1}) .$$

Since the functions $p(\cdot, \theta^m) - \mu^*(\theta^m)q(\cdot, \theta^m)$ and $p(\cdot, \theta^{m+1}) - \mu^*(\theta^{m+1})q(\cdot, \theta^{m+1})$ are both linear over $[\hat{v}^m, 1]$, it follows that $-\mu^*(\theta^m) \geq -\mu^*(\theta^{m+1})$ and $\hat{v}^m \geq \hat{v}^{m+1}$. Since the function μ^* is bounded (Lemma 1), we know that $\lim_{m \rightarrow \infty} -\mu^*(\theta^m) = -\underline{\mu}$.

We consider three sequences: (i) $\{(v^m, \lambda(v^m))\}_{m \in \mathbb{Z}^+}$, (ii) $\{(\theta^m + c, u(\theta^m))\}_{m \in \mathbb{Z}^+}$, and (iii) $\{(\hat{v}^m, \lambda(\hat{v}^m))\}_{m \in \mathbb{Z}^+}$. Notice that for any $m \in \mathbb{Z}^+$, three points $(v^m, \lambda(v^m))$, $(\theta^m + c, u(\theta^m))$, $(\hat{v}^m, \lambda(\hat{v}^m))$ lie in a common linear function $-\mu^*(\theta^m)q(\cdot, \theta^m) + p(\cdot, \theta^m)$ over $[\theta^m, 1 - c]$. Let (x_a, y_a) denote the limit point of the sequence $\{(v^m, \lambda(v^m))\}_{m \in \mathbb{Z}^+}$. Let (x_b, y_b) denote the limit point of the sequence $\{(\theta^m + c, u(\theta^m))\}_{m \in \mathbb{Z}^+}$. Let (x_c, y_c) denote the limit point of the sequence $\{(\hat{v}^m, \lambda(\hat{v}^m))\}_{m \in \mathbb{Z}^+}$. We can observe that the function $-\mu^*(\theta)q(\cdot, \theta) + p(\cdot, \theta)$ crosses the point (x_b, y_b) . It follows that $u'(\theta) \geq -\underline{\mu}$.

Next we aim to prove $u'(\theta) = -\underline{\mu}$ through contradiction. We assume that $u'(\theta) > -\underline{\mu}$. We divide the proof into two cases.

Subcase 2.1: Points (x_b, y_b) and (x_c, y_c) do not overlap. We show that $\frac{y_c - y_b}{x_c - x_b} = -\underline{\mu}$.

$$\begin{aligned} \frac{y_c - y_b}{x_c - x_b} &= \lim_{m \rightarrow \infty} \frac{\lambda^*(\hat{v}^m) - u(\theta^m)}{\hat{v}^m - (\theta^m + c)} \\ &= \lim_{m \rightarrow \infty} \frac{-\mu^*(\theta^m)[\hat{v}^m - (\theta^m + c)]}{\hat{v}^m - (\theta^m + c)} \\ &= \lim_{m \rightarrow \infty} -\mu^*(\theta^m) = -\underline{\mu}. \end{aligned}$$

In a similar manner, we can also prove that $\frac{y_b - y_a}{x_b - x_a} = -\underline{\mu}$. Thus, we have proved that points (x_a, y_a) , (x_b, y_b) , (x_c, y_c) lie in a common linear function. Next we consider an index $\theta + \epsilon$ for sufficiently small $\epsilon > 0$. Since the function λ^* is continuous, we know that

$$-\mu^*(\theta)q(x_c, \theta) + p(x_c, \theta) \geq -\mu^*(\theta + \epsilon)q(x_c, \theta + \epsilon) + p(x_c, \theta + \epsilon),$$

which further implies that $-\mu^*(\theta + \epsilon) < -\underline{\mu}$. Finally, we know that

$$-\mu^*(\theta + \epsilon)q(v, \theta + \epsilon) + p(v, \theta + \epsilon) = u(v) + \mu^*(\theta + \epsilon) \cdot c > u(v) + \underline{\mu} \cdot c = \lambda(v),$$

which forms a contradiction to Lemma 2. Thus, we have proved that $u'(\theta) = -\underline{\mu}$.

Subcase 2.2: Points (x_b, y_b) and (x_c, y_c) overlap. For any $m \in \mathbb{Z}^+$, it holds $\lambda(\hat{v}^m) \geq u(\hat{v}^m - c)$. The fact that points (x_b, y_b) and (x_c, y_c) overlap implies that $\lambda^*(\theta + c) = u(\theta)$. Further,

$$\frac{\lambda(\hat{v}^m) - \lambda(\theta + c)}{\hat{v}^m - \theta - c} \geq \frac{u(\hat{v}^m - c) - u(\theta)}{\hat{v}^m - \theta - c}.$$

Based on the graph illustrations of these functions, we know $-\mu^*(\theta^m) \geq \frac{\lambda(\hat{v}^m) - \lambda(\theta + c)}{\hat{v}^m - \theta - c}$. Thus, we have that

$$-\mu^*(\theta^m) \geq \frac{u(\hat{v}^m - c) - u(\theta)}{\hat{v}^m - \theta - c}.$$

Passing m to infinity, we achieve that

$$-\underline{\mu} = \lim_{m \rightarrow \infty} -\mu^*(\theta^m) \geq u'(\theta),$$

which forms a contradiction with the former assumption.

By combining these two cases, we have proved that $-\underline{\mu} = u'(\theta)$. We know that $\frac{y_b - y_a}{x_b - x_a} = -\underline{\mu}$. Thus, we have that

$$u'(\theta) = -\underline{\mu} = \frac{y_b - y_a}{x_b - x_a} = \frac{u(\theta) - \lambda^*(v)}{\theta + c - v},$$

which implies that $\lambda^*(v) = u(\theta) + u'(\theta)(v - \theta - c)$.

Subcase 2.3: Otherwise. For any $\epsilon > 0$, we define two sets

$$A \triangleq \{(v', \theta') | v' \geq \theta', v' \in [v - \epsilon, v]\}, \quad B \triangleq \{(v', \theta') | \theta' \geq v, v' \in [v - \epsilon, v]\}.$$

In this case, there exists an $\epsilon > 0$ such that regions A and B are both empty. In other words, for any value $v' \in (v - \epsilon, v)$, there exists $\theta(v') \in (v', v)$ such that $(v', g(v')) \in \text{supp}(G^*)$. Next, we define function $h(v') = -\mu^*(\theta(v'))$ for any $v' \in (v - \epsilon, v)$. For any $v_1, v_2 \in (v - \epsilon, v)$ with $v_1 < v_2$, it follows that

$$h(v_1) = \min_{\theta' \in [v_1, 1-c]} -\mu^*(\theta'), \quad h(v_2) = \min_{\theta' \in [v_2, 1-c]} -\mu^*(\theta').$$

Thus we know that $h(v_2) \geq h(v_1)$, which implies that the function h is weakly increasing over $(v - \epsilon, v)$. Next we are going to prove that the function h is also continuous over $(v - \epsilon, v)$.

First, we show the function h is right-continuous over $(v - \epsilon, v)$. For any $w \in (v', g(v'))$, it holds that $h(w) = h(v')$. Since the prior has positive density everywhere, we know that the function h is right-continuous over $(v - \epsilon, v)$.

Then, we show the function h is left-continuous over $(v - \epsilon, v)$. We prove it through contradiction. We assume that the function h is not left-continuous everywhere over $(v - \epsilon, v)$. For any tuple $(v', g(v')) \in \text{supp}(G^*)$, we define sets

$$R \triangleq \{(v'', \theta'') | \theta'' \in (v', v), v'' \in [0, v']\}, \quad S \triangleq \{(v'', \theta'') | v'' \in (v - \epsilon, v'), \theta'' \in (v', v)\}.$$

Under this case, the region R must be empty. Besides, for any value $v'' \in (v - \epsilon, v')$, if $(v'', \theta'') \in \text{supp}(G^*)$, then it must hold that $(v'', \theta'') \in S$. Since h is increasing, we know that $\lim_{w \rightarrow v'^-} h(w) = \hat{h}$. Since $\lim_{w \rightarrow v'^-} \lambda^*(w) = \lambda^*(v')$, we know that $u(v') - c\hat{h} = u(v') - ch(v')$. Finally, we achieve that $\hat{h} = h(v')$, which forms a contradiction with the assumption. Till now, we have proved that the function h is continuous over $(v - \epsilon, v)$.

Next we aim to prove that h is flat over $(v - \epsilon, v)$. We assume that the image of the function h includes some open interval (y_1, y_2) , thus the image is uncountable. Then, we define a set

$$T(y) \triangleq \{x | h(x) = y\}.$$

It is obvious that $T(y)$ is non-empty, and we let $z \in T(y)$. We know that $T(y)$ includes $(z, g(z))$. Therefore, we know that for any given y , its corresponding set $T(y)$ will consist of mutually disjoint intervals $(z, g(z))$. If we arbitrarily select a rational number from each such interval, then we would have obtained uncountably many rational numbers within the interval $(v - \epsilon, v)$. However, we know that the number of rational numbers within any interval is countable. This leads to a contradiction.

To sum up, we have proved that, under this case, there exists an $\delta > 0$, such that $-\mu^*(\theta') = u'(\theta)$ for any $\theta' \in (\theta - \delta, \theta]$. Thus, we directly have that

$$\lambda^*(v) = u'(\theta)q(v, \theta) + p(v, \theta).$$

This completes the proof. $\square$

Lemma 15 (Case 3.3 in Algorithm 1). *Given G^* and (λ^*, μ^*) are both optimal to the primal and dual problems, if there exists (v, θ) such that $v \in [\underline{v}, 1]$, $v \neq \theta + c$, and $v > v' = \inf \text{supp}(G_{\cdot|\theta})$, then we have*

$$\lambda^*(v) = \frac{p(v, \theta)q(v', \theta) - p(v', \theta)q(v, \theta) + \lambda^*(v')q(v, \theta)}{q(v', \theta)}.$$

Proof of Lemma 15. Since $(v, \theta) \in \text{supp}(G)$, by Lemma 2, we know that $\lambda^*(v) = -\mu^*(\theta)q(v, \theta) + p(v, \theta)$. Since the functions λ^* , p , and q are all continuous (Lemma 7), it holds that $\lambda^*(v') = -\mu^*(\theta)q(v', \theta) + p(v', \theta)$. By combining these two equations, we have

$$\lambda^*(v) = \frac{p(v, \theta)q(v', \theta) - p(v', \theta)q(v, \theta) + \lambda^*(v')q(v, \theta)}{q(v', \theta)}.$$

This completes the proof. $\square$

Lemma 16 (Algorithm 3). *Given G^* and (λ^*, μ^*) are both optimal to the primal and dual problems, for any $\theta \in \text{supp}(G_\theta^*)$ and any two distinct $v_1, v_2 \neq \theta + c$ such that $(v_1, \theta), (v_2, \theta) \in \text{supp}(G^*)$, we have that*

$$\mu^*(\theta) = \frac{p(v_1, \theta) - \lambda^*(v_1)}{q(v_1, \theta)} = \frac{p(v_2, \theta) - \lambda^*(v_2)}{q(v_2, \theta)} .$$

Proof of Lemma 16. It holds directly by Lemma 2. $\square$

Lemma 17 (Algorithm 3). *Given G^* and (λ^*, μ^*) are both optimal to the primal and dual problems, for any $\theta \in [-c, 1 - c]$, it holds that*

$$\max_{v \in (\theta+c, 1]} \frac{u(\theta) - \lambda^*(v)}{v - \theta - c} \leq \mu^*(\theta) \leq \min \left\{ \min_{v \in [0, \theta)} \frac{\lambda^*(v) - u(v)}{c}, \min_{v \in [\theta, \theta+c)} \frac{u(\theta) - \lambda^*(v)}{v - \theta - c} \right\} .$$

Proof of Lemma 17. By Constraint (1) and Lemma 2, we directly have that

$$\max_{v \in (\theta+c, 1]} \frac{u(\theta) - \lambda^*(v)}{v - \theta - c} \leq \mu^*(\theta) \leq \min \left\{ \min_{v \in [0, \theta)} \frac{\lambda^*(v) - u(v)}{c}, \min_{v \in [\theta, \theta+c)} \frac{u(\theta) - \lambda^*(v)}{v - \theta - c} \right\} .$$

This completes the proof. $\square$

With the above lemmas, finally, we can provide the whole proof of Theorem 1.

Proof of Theorem 1. This theorem directly holds by Lemmas 3 and 8 to 17. $\square$

B.6 Equilibrium Characterization for Convex Priors with Low Cost

Theorem 4. *If all senders share the same prior F that is (weakly) convex, with all values in its support at least c , then each sender fully revealing his value is the unique symmetric equilibrium.*

Proof of Theorem 4. When all senders adopt the full-revelation strategy, it is easy to check that the corresponding strategy profile induced is a symmetric equilibrium. We now show that this is the unique symmetric equilibrium.

Since we consider symmetric equilibria, we omit the subscripts for senders. For the sake of contradiction, suppose G^* is another symmetric equilibrium strategy that does not fully reveal the value. Let K denote the distribution of the amortized value induced by the 2-D distribution G^* . By the MPC condition, we know that K shifted by c is a MPC of F , i.e.,

$$\int_{-c}^t K(x) \, dx \leq \int_0^{t+c} F(x) \, dx, \quad \forall t \in [-c, 1 - c] .$$

Besides, we have $K(-c) = F(0) = 0$, $K(1 - c) = F(1) = 1$, and

$$\int_{-c}^{1-c} K(x) \, dx = \int_0^1 F(x) \, dx .$$

Let $u(v)$ denote the contribution of value $v \in [c, 1]$ to the sender's expected utility. Then, we can rewrite the expected utility of the sender as $\int_c^1 u(v) f(v) \, dv$. Under the fact that strategy G^* forms a symmetric equilibrium, function K must be continuous over $[-c, 1 - c]$; otherwise, each sender has a profitable deviation by spreading the index where the mass is. Under this fact, we have the corresponding function $\phi(x) = K^{N-1}(x)$ for any $x \in [-c, 1 - c]$. Each sender can always choose to send no information, thereby obtaining an expected utility of $\int_c^1 K^{N-1}(v - c) f(v) \, dv$.

Therefore, each sender is guaranteed to achieve at least an expected utility of $\int_c^1 K^{N-1}(v-c)f(v) dv$, that is

$$\begin{aligned} \int_c^1 u(v)f(v) dv &\geq \int_c^1 K^{N-1}(v-c)f(v) dv = \int_c^1 f(v) d\left(\int_{-c}^{v-c} K^{N-1}(t) dt\right) \\ &= \left[f(v) \int_{-c}^{v-c} K^{N-1}(t) dt\right] \Big|_{v=c}^1 - \int_c^1 \int_{-c}^{v-c} K^{N-1}(t) dt df(v) \\ &= \left[f(v) \int_c^v F^{N-1}(t) dt\right] \Big|_{v=c}^1 - \int_c^1 \int_{-c}^{v-c} K^{N-1}(t) dt df(v). \end{aligned}$$

Since $\int_{-c}^{v-c} K^{N-1}(t) dt \leq \int_c^v F^{N-1}(t) dt$ for any $v \in [c, 1]$ and the density of the prior is non-decreasing over $[c, 1]$, we have

$$\begin{aligned} \int_c^1 u(v)f(v) dv &\geq \left[f(v) \int_c^v F^{N-1}(t) dt\right] \Big|_{v=c}^1 - \int_c^1 \int_c^v F^{N-1}(t) dt df(v) \\ &= \int_c^1 F^{N-1}(v) dF(v) = \frac{1}{N}. \end{aligned}$$

In any symmetric equilibrium of the N -senders game, each sender can obtain at most an expected utility of $1/N$. Combining these two facts, we obtain that $\int_c^1 u(v)f(v) dv = 1/N$. Therefore, all the inequalities above must hold with equality, where we have

$$\int_{-c}^t K(x) dx = \int_0^{t+c} F(x) dx, \quad \forall t \in [-c, 1-c].$$

This fact is equivalent to the fact that the symmetric equilibrium strategy G^* is indeed induced by the full-revelation strategy, which forms a contradiction with the assumption. Thus, we have proved that the symmetric equilibrium induced by the full-revelation strategy is the unique symmetric equilibrium. $\square$

For an intuition that full revelation constitutes an equilibrium, observe that, for each sender, the interim utility $u(\kappa) = F^{N-1}(\kappa+c)$ is convex in this scenario; any pooling of values decreases the expected utility by Jensen's inequality. Intuitively, fully revealing the values is a best response.

The uniqueness of this symmetric equilibrium is nontrivial. As a starting point, note that the senders' game is constant-sum for such a small c : since all values are at least c , the index of each sender is always non-negative, regardless of the signaling; therefore, the buyer never quits without selecting a seller. Thus, in any symmetric equilibrium, each sender must share a common utility of $1/N$. We show that if all senders except i use a common strategy which does not always fully reveal the values, sender i can get a utility strictly higher than $1/N$ by fully revealing v_i . The uniqueness of the symmetric equilibrium follows.

Note that the market achieves maximal efficiency under this equilibrium: the receiver fully learns the relevant information from the senders' signaling and therefore always selects the best box, attaining first-best revenue.

B.7 Proof of Theorem 2

Proof of Theorem 2. Denote the shifted prior by $\tilde{F}(x) := F(x+c)$ for $x \in [-c, 1-c]$ (with density $\tilde{f}$). By the MPC condition, we know that the amortized value distribution of any feasible strategy forms an MPC of the function $\tilde{F}$ over $[-c, 1-c]$. We divide this proof into three steps. Here we provide the proof of the non-degenerate case where $\theta'_1 = \theta_2 < 1-c$.

Step 1: There uniquely exist $\theta_1 \in (-c, 0)$ and $\theta_2 \in (0, 1 - c]$ such that $\frac{\tilde{F}(\theta_1)}{-\theta_1} = \frac{\tilde{F}(\theta_2) - \tilde{F}(\theta_1)}{\theta_2}$, and $K \in \text{MPC}(\tilde{F})$ over $[-c, 1 - c]$. For any $\theta_1 \in (-c, 0)$, we can construct the following linear function:

$$L_{\theta_1}(\theta) = \frac{\tilde{F}(\theta_1)}{-\theta_1}(\theta - \theta_1), \quad \forall \theta \in [\theta_1, 1 - c].$$

There exists an $\epsilon_1 \in (0, c)$ such that for any $\theta_1 \in (-c, -c + \epsilon)$, it holds that $L_{\theta_1}(\theta) \leq \tilde{F}(\theta)$ for any $\theta \in [\theta_1, 1 - c]$. There exists an $\epsilon_2 \in (0, c)$ such that for any $\theta_1 \in (-\epsilon_2, 0)$, the function L_{θ_1} has at least one intersection point with function $\tilde{F}$. Since $\tilde{F}(\theta_1)$ and $\tilde{F}(\theta_1)/-\theta_1$ are both strictly increasing in θ_1 over $(-c, 0)$, we know that there uniquely exists a $\underline{\theta}_1 \in (-c, 0)$ such that $L_{\underline{\theta}_1}(\theta) \leq \tilde{F}(\theta)$ for any $\theta \in [\underline{\theta}_1, 1 - c]$, and function $L_{\underline{\theta}_1}$ and function $\tilde{F}$ are tangent at some points. Furthermore, there uniquely exists a $\bar{\theta}_1 \in (\underline{\theta}_1, 0)$ such that $L_{\bar{\theta}_1}(1 - c) = \tilde{F}(1 - c)$. Thus for any $\theta_1 \in (\underline{\theta}_1, 0)$, function L_{θ_1} and function $\tilde{F}$ have at least two intersection points within the interval $[\theta_1, 1 - c]$. For any $\theta_1 \in (\underline{\theta}_1, 0)$, we define $\theta_2 = \max\{\theta \in [0, 1 - c] | L_{\theta_1}(\theta) = \tilde{F}(\theta)\}$ and $\theta_3 = \min\{\theta \in [0, 1 - c] | L_{\theta_1}(\theta) = \tilde{F}(\theta)\}$. Using the three new parameters, we can split the integral difference from θ_1 to θ_2 into two parts:

$$\begin{aligned} & \int_{\theta_1}^{\theta_2} \tilde{F}(\theta) d\theta - \int_{\theta_1}^{\theta_2} L_{\theta_1}(\theta) d\theta \\ &= \left(\int_{\theta_3}^{\theta_2} \tilde{F}(\theta) d\theta - \int_{\theta_3}^{\theta_2} L_{\theta_1}(\theta) d\theta \right) - \left(\int_{\theta_1}^{\theta_3} L_{\theta_1}(\theta) d\theta - \int_{\theta_1}^{\theta_3} \tilde{F}(\theta) d\theta \right). \end{aligned}$$

There exists an $\epsilon_3 \in (0, \bar{\theta}_1 - \underline{\theta}_1)$ such that $\int_{\theta_1}^{\theta_2} \tilde{F}(\theta) d\theta - \int_{\theta_1}^{\theta_2} L_{\theta_1}(\theta) d\theta < 0$ holds for any $\theta_1 \in (\underline{\theta}_1, \underline{\theta}_1 + \epsilon_3)$. There exists an $\epsilon_4 \in (0, \bar{\theta}_1 - \underline{\theta}_1)$ such that $\int_{\theta_1}^{\theta_2} \tilde{F}(\theta) d\theta - \int_{\theta_1}^{\theta_2} L_{\theta_1}(\theta) d\theta > 0$ holds for any $\theta_1 \in (\underline{\theta}_1, \underline{\theta}_1 + \epsilon_4)$. Besides, the first part $\int_{\theta_3}^{\theta_2} \tilde{F}(\theta) d\theta - \int_{\theta_3}^{\theta_2} L_{\theta_1}(\theta) d\theta$ is decreasing at θ_1 over $(\underline{\theta}_1, 0)$, while the second part $\int_{\theta_1}^{\theta_3} L_{\theta_1}(\theta) d\theta - \int_{\theta_1}^{\theta_3} \tilde{F}(\theta) d\theta$ is increasing at θ_1 over $(\underline{\theta}_1, 0)$. So we know that there uniquely exists a $\theta_1 \in (\underline{\theta}_1, \bar{\theta}_1)$ such that $\int_{\theta_1}^{\theta_2} \tilde{F}(\theta) d\theta - \int_{\theta_1}^{\theta_2} L_{\theta_1}(\theta) d\theta = 0$. Finally, we prove that such a pair of parameters (θ_1, θ_2) exists and is unique.

Step 2: Given θ_1 and θ_2 determined in Step 1, there exists a 2-D distribution that induces the hinge-shaped amortized value distribution. We prove this by a construction. We divide this part of the proof into two cases, based on the shape of K , the amortized value distribution.

Case 1: $K(\theta'_1) < 1$. In this case, $\theta_2 = \theta'_1 < 1 - c$. Values in the $[0, \theta_1 + c]$ and $[\theta_2 - c, 1]$ are fully revealed, so that K and $\tilde{F}$ coincide on these intervals. We are to pool values in $[\theta_1 + c, \theta_2 + c]$ to produce amortized values uniformly distributed over $[0, \theta_2]$. Next, we show how to make such poolings.

Based on the structure of the function K , the slope over $[0, \theta_2]$ is denoted by $\rho = \frac{\tilde{F}(\theta_2) - \tilde{F}(\theta_1)}{\theta_2}$. Since $K \in \text{MPC}(\tilde{F})$, it holds that $\tilde{f}(0) < \rho$ and $\tilde{f}(\theta_2) > \rho$. Besides, the density $\tilde{f}$ is increasing over $[0, \theta_2]$. So there exists a row $\theta^* \in (0, \theta_2)$ such that $\tilde{f}(\theta^*) = K(\theta^*) = \rho$. We need to ensure the index equation is preserved while pooling the values from $[\theta_1 + c, \theta_2 + c]$ into indices $[0, \theta_2]$, with the density of each row equal to ρ . We divide the density of each value $v \in [\theta^* + c, \theta_2 + c]$ into two parts: we keep a portion of size ρ unpooled to individually form an index $v - c$, while the remaining density will be pooled together with lower values. In details, first we let $\text{supp}(G_{|\theta}) = \{\theta + c\}$ and $g(\theta + c, \theta) = \rho$ for any $\theta \in (\theta^*, \theta_2)$. Then, we show how to pool the values from $[\theta_1 + c, \theta^* + c]$ and $[\theta^* + c, \theta_2 + c]$ together to form indices over $[0, \theta^*]$. We define functions $\alpha : [0, \theta^*] \rightarrow [0, 1]$ and $\beta : [0, \theta^*] \rightarrow [0, 1]$. We assume that the support of the conditional distribution G_θ is $\{\alpha(\theta), \beta(\theta)\}$ for each $\theta \in [0, \theta^*]$. We give the initial values of these two functions as $\alpha(0) = \theta_1 + c$ and $\beta(0) = \theta_2 + c$. Given the initial conditions for the functions α and β , we now only need to compute their derivatives to fully characterize these two functions.

We consider the following two cases to show how to specifically compute the derivatives of these two functions.

case-1.1: $\alpha(\theta) > \theta$. In this case, we have these two equations:

$$\frac{f(\alpha(\theta))}{\alpha'(\theta)} + \frac{f(\beta(\theta)) - \rho}{-\beta'(\theta)} = \rho, \quad (\alpha(\theta) - \theta) \frac{f(\alpha(\theta))}{\alpha'(\theta)\rho} + (\beta(\theta) - \theta) \frac{f(\beta(\theta)) - \rho}{-\beta'(\theta)\rho} = c.$$

The first equation means that the sum of densities at points $(\alpha(\theta), \theta)$ and $(\beta(\theta), \theta)$ is equal to ρ . The second equation is exactly the index equation of row θ . Both equations are ordinary differential equations. By combining these two equations, we achieve that:

$$\alpha'(\theta) = \frac{(\beta(\theta) - \alpha(\theta))f(\alpha(\theta))}{\rho(\beta(\theta) - c - \theta)}, \quad \beta'(\theta) = \frac{(\beta(\theta) - \alpha(\theta))(\rho - f(\beta(\theta)))}{\rho(\theta + c - \alpha(\theta))}.$$

case-1.2: $\alpha(\theta) \leq \theta$. In this case, we have these two equations:

$$f(\alpha(\theta)) + \frac{f(\beta(\theta)) - \rho}{-\beta'(\theta)} = \rho, \quad (\beta(\theta) - \theta) \frac{f(\beta(\theta)) - \rho}{-\beta'(\theta)} = c \left(\frac{f(\alpha(\theta))}{\alpha'(\theta)} + \frac{f(\beta(\theta)) - \rho}{-\beta'(\theta)} \right).$$

The first equation means that the sum of densities at points $(\alpha(\theta), \theta)$ and $(\beta(\theta), \theta)$ is equal to ρ . The second equation is exactly the index equation of row θ . By combining these two equations, we achieve that:

$$\alpha'(\theta) = \frac{-cf(\alpha(\theta))}{(f(\alpha(\theta)) - \rho)(\beta(\theta) - \theta - c)}, \quad \beta'(\theta) = \frac{f(\beta(\theta)) - \rho}{f(\alpha(\theta)) - \rho}.$$

In both subcases, for each $\theta \in [0, \theta^*]$, since the value $\alpha(\theta)$ and the value $\beta(\theta)$ are pooled together to form the index θ , we have $\alpha(\theta) \leq \theta + c \leq \beta(\theta)$. Since the prior density f is increasing, we achieve that $\alpha'(\theta) \geq 0$ if $\alpha(\theta) \leq \theta^* + c$, and $\beta'(\theta) \leq 0$ if $\beta(\theta) \geq \theta^* + c$. In addition to the initial values of both functions, there always exists some $\theta' \in [0, \theta^*]$ such that the function α increases over the interval $[0, \theta']$, while the function β decreases over the same interval. We define the maximal such θ' as $\tilde{\theta}$, i.e., $\tilde{\theta} \triangleq \sup\{\theta \in [0, \theta^*] \mid \alpha'(\theta) \geq 0 \text{ \& } \beta'(\theta) \leq 0\}$. We next show $\tilde{\theta} = \theta^*$ via contradiction. We assume that $\tilde{\theta} < \theta^*$, then it must hold that $\beta(\tilde{\theta}) = \theta^* + c$ and $\alpha(\tilde{\theta}) = \tilde{\theta} + c$. Since the function K constructed in Step-1 forms an MPC of the function $\tilde{F}$ over $[\theta_1, \theta_2]$, we have

$$F(\theta_2 + c) - F(\theta_1 + c) = \rho \cdot \theta_2.$$

Besides, based on the construction above, we have

$$\rho \cdot (\theta_2 - \theta^* + \tilde{\theta}) = F(\theta_2 + c) - F(\theta^* + c) + F(\tilde{\theta} + c) - F(\theta_1 + c).$$

Combining these two equations, we achieve that

$$\rho \cdot (\theta^* - \tilde{\theta}) = F(\theta^* + c) - F(\tilde{\theta} + c),$$

which forms a contradiction since $F(\theta^* + c) - F(\tilde{\theta} + c) < \rho \cdot (\theta^* - \tilde{\theta})$. Thus, we have proved that $\tilde{\theta} = \theta^*$. Under this fact, we complete the proof by showing that $\alpha(\theta^*) = \beta(\theta^*) = \theta^* + c$. We also prove it through contradiction. Suppose $\alpha(\theta^*) = a < \theta^* + c < b = \beta(\theta^*)$. Since the constructed function K forms an MPC of the function $\tilde{F}$ over $[\theta_1, \theta_2]$, it holds that

$$F(\theta_2 + c) - F(\theta_1 + c) = \rho \cdot \theta_2.$$

Besides, based on the construction above, we have

$$\rho \cdot \theta_2 = F(\theta_2 + c) - F(b) + F(a) - F(\theta_1 + c).$$

These two equations form a contradiction, which makes the assumption invalid. Till now, we have proved that the method we provided above indeed can construct such 2-D distribution that induces the amortized value distribution defined in Theorem 2.

Case 2: There exists $\epsilon > 0$ such that the amortized value distribution K is constant over $(1 - c - \epsilon, 1 - c)$. The construction and proof follow a similar (and much simpler) approach to Case 1, so we omit them here.

Step 3: Equilibrium verification. Based on the techniques in Section 3, we can verify that the strategy constructed above indeed constitutes a best response to the interim utility function u . Using the Algorithms 1 and 3, we can construct such λ and μ :

$$\lambda(v) = \begin{cases} K(v - c) & \text{if } v \in [0, \theta_1 + c) \cup [c, 1] , \\ \rho(v - c) + F(\theta_1 + c) & \text{if } v \in [\theta_1 + c, c) , \end{cases}$$

and

$$\mu(\theta) = \begin{cases} \lambda'(\theta + c) & \text{if } \theta \in [0, \theta_1 + c) \cup (\theta_1 + c, \theta_2 + c) \cup (\theta_2 + c, 1] , \\ \rho & \text{if } \theta \in \{\theta_1 + c, \theta_2 + c\} . \end{cases}$$

Based on the formulas of λ and μ , we directly have that for any $(v, \theta) \in \text{supp}(G)$:

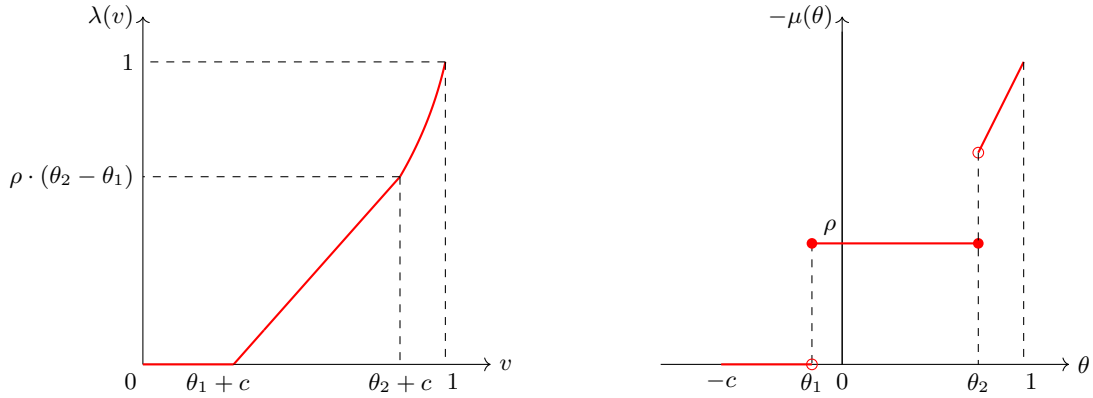


Fig. 6. The dual variables constructed by the Recursive Reduction Procedure, based on the symmetric equilibrium constructed in Theorem 2

$$\lambda(v) = p(v, \theta) - \mu(\theta)q(v, \theta) ,$$

which implies that G and (λ, μ) satisfy the Complementary Slackness conditions in Lemma 2. Besides, it also can be verified that such λ and μ are feasible to the corresponding dual problem, that is:

$$\lambda(v) = \max_{\theta \in [-c, 1-c]} p(v, \theta) - \mu(\theta)q(v, \theta) , \quad \forall v \in [0, 1] ,$$

which makes (λ, μ) feasible to the corresponding dual problem. Combining these two results and by Theorem 1, we know that such G indeed constitutes a best response to the interim utility function u . $\square$

C Proofs in Section 4

Proof of Theorem 3. We proceed with the proof in the following three steps.

Step 1: Discretizing the original prior F . Fix a value-grid size $\delta \in (0, 1)$. Let $M \triangleq \lceil 1/\delta \rceil$ be the number of intervals and partition $[0, 1]$ into intervals

$$I_r \triangleq \begin{cases} [\frac{r-1}{M}, \frac{r}{M}) & \text{if } r \in [M-1] , \\ [\frac{M-1}{M}, 1] & \text{if } r = M . \end{cases}$$

For each interval I_r , let $\hat{v}_r$ be its midpoint $\hat{v}_r \triangleq \frac{2r-1}{2M}$ for each $r \in [M]$. Then, we move all probability mass of the original prior F inside each interval I_r to the representative value $\hat{v}_r$.⁹

$$F_\delta \triangleq \sum_{r=1}^M F(I_r) \cdot \delta_{\hat{v}_r} ,$$

where $\delta_{\hat{v}_r}$ denotes the Dirac measure at $\hat{v}_r$.

Under the discretized prior F_δ , we obtain the following auxiliary best-response problem:

$$\max_{G_\delta \in \mathcal{G}(F_\delta, c)} U(G_\delta) = \mathbb{E}_{(v, \theta) \sim G_\delta} [u(\min\{v, \theta\})] . \quad (\mathcal{P}_{\text{BR}}^\delta)$$

Since the feasible space $\mathcal{G}(F_\delta, c)$ is convex and weakly compact, and $u(\kappa)$ is continuous, problem $(\mathcal{P}_{\text{BR}}^\delta)$ still admits an optimal solution. Let G_δ^* be an optimal solution to the problem $(\mathcal{P}_{\text{BR}}^\delta)$, and denote its optimal value by $\text{OPT}_\delta \triangleq U(G_\delta^*)$. Notice that $(\mathcal{P}_{\text{BR}})$ and $(\mathcal{P}_{\text{BR}}^\delta)$ are not nested optimization problems. Indeed, their feasible sets are $\mathcal{G}(F, c)$ and $\mathcal{G}(F_\delta, c)$, respectively, and neither feasible set contains the other. Therefore, there is no direct monotonicity relation between OPT and OPT_δ . Instead, we prove a stability bound in two steps:

$$|\text{OPT}_\delta - \text{OPT}| \leq L\delta .$$

Step 1.1: From the original prior to the discretized prior. Starting from the optimal solution $G^* \in \mathcal{G}(F, c)$ of the original problem $(\mathcal{P}_{\text{BR}})$, we construct a feasible solution $G_\delta \in \mathcal{G}(F_\delta, c)$ for the discretized-prior problem $(\mathcal{P}_{\text{BR}}^\delta)$, and prove that the construction loss is at most $L\delta$. For G_θ^* -almost every row θ , we construct a discretized posterior by moving all mass in each value interval I_r to its midpoint $\hat{v}_r$. Specifically, write $m_r(\theta) \triangleq G_{v|\theta}^*(I_r)$ for all $r \in [M]$. Then we define the discretized posterior on the value grid by

$$\hat{G}_{v|\theta} \triangleq \sum_{r=1}^M m_r(\theta) \cdot \delta_{\hat{v}_r} .$$

Let $\theta_\delta(\theta)$ be the index of this discretized posterior, i.e., the unique solution to

$$\int_0^1 ((v - \theta_\delta(\theta))_+ - c) \, d\hat{G}_{v|\theta}(v) = 0 .$$

Thus, each original row θ is transformed as follows: value mass in every interval I_r is moved horizontally to $\hat{v}_r$, and the whole row with newly moved masses at each midpoint is then moved vertically from θ to the recomputed index $\theta_\delta(\theta)$. Formally, define G_δ by

$$G_\delta(A \times B) \triangleq \int_{-c}^{1-c} \mathbf{1}\{\theta_\delta(\theta) \in B\} \hat{G}_{v|\theta}(A) \, dG_\theta^*(\theta)$$

for all measurable $A \subseteq [0, 1]$ and $B \subseteq [-c, 1-c]$.

⁹ The specific representative point inside each interval is immaterial for the approximation argument; we use the midpoint only for convenience.

We first show that the vertical and horizontal movements are both small for every probability mass. To see this, for every $v \in I_r$, we have $|v - \hat{v}_r| \leq \delta$. Hence, for every row θ ,

$$(\hat{v}_r - (\theta + \delta))_+ \leq (v - \theta)_+ , \quad (\hat{v}_r - (\theta - \delta))_+ \geq (v - \theta)_+ , \quad \forall v \in I_r .$$

Integrating over all intervals with respect to $G_{v|\theta}^*$ gives

$$\int_0^1 (\hat{v} - (\theta + \delta))_+ d\hat{G}_{v|\theta}(\hat{v}) \leq c , \quad \int_0^1 (\hat{v} - (\theta - \delta))_+ d\hat{G}_{v|\theta}(\hat{v}) \geq c .$$

By monotonicity of the index equation in the index variable,

$$|\theta_\delta(\theta) - \theta| \leq \delta \quad \text{for } G_\theta^* \text{-almost every } \theta .$$

It remains to compare payoffs. Recall that, for each original row θ , all mass in interval I_r is moved to the representative value $\hat{v}_r$, and the row index is recomputed as $\theta_\delta(\theta)$. For every $v \in I_r$, we have $|v - \hat{v}_r| \leq \delta$, and from the index-stability argument above, $|\theta - \theta_\delta(\theta)| \leq \delta$ for G_θ^* -almost every θ . Therefore, for every $v \in I_r$ and G_θ^* -almost every θ ,

$$|\min\{\hat{v}_r, \theta_\delta(\theta)\} - \min\{v, \theta\}| \leq \delta .$$

Since the interim utility $u(\kappa)$ is L -Lipschitz continuous, this implies

$$u(\min\{\hat{v}_r, \theta_\delta(\theta)\}) \geq u(\min\{v, \theta\}) - L\delta .$$

Applying the pointwise payoff bound on each interval I_r , we obtain

$$\begin{aligned} U(G_\delta) &= \int_{-c}^{1-c} \sum_{r=1}^M u(\min\{\hat{v}_r, \theta_\delta(\theta)\}) G_{v|\theta}^*(I_r) dG_\theta^*(\theta) \\ &\geq \int_{-c}^{1-c} \sum_{r=1}^M \int_{I_r} [u(\min\{v, \theta\}) - L\delta] dG_{v|\theta}^*(v) dG_\theta^*(\theta) \\ &= U(G^*) - L\delta = \text{OPT} - L\delta . \end{aligned}$$

Since G_δ^* is optimal for $(\mathcal{P}_{\text{BR}}^\delta)$, we conclude that

$$\text{OPT}_\delta \geq U(G_\delta) \geq \text{OPT} - L\delta .$$

Step 1.2: From the discretized prior back to the original prior. Conversely, starting from the optimal solution $G_\delta^* \in \mathcal{G}(F_\delta, c)$ of the discretized-prior problem $(\mathcal{P}_{\text{BR}}^\delta)$, we construct a feasible solution $G \in \mathcal{G}(F, c)$ for the original problem $(\mathcal{P}_{\text{BR}})$.

For $G_{\delta, \theta}^*$ -almost every row θ , write $m_r(\theta) \triangleq G_{\delta, v|\theta}^*(\{\hat{v}_r\})$ for all $\forall r \in [M]$. We lift the discretized posterior by spreading the mass at each midpoint $\hat{v}_r$ back to the interval I_r according to the conditional distribution of F on I_r . Specifically, define a posterior distribution $\tilde{G}_{v|\theta}$ on $[0, 1]$ by

$$\tilde{G}_{v|\theta}(A) \triangleq \sum_{r=1}^M m_r(\theta) \frac{F(A \cap I_r)}{F(I_r)}$$

for every measurable set $A \subseteq [0, 1]$. It holds $F(I_r) > 0$ for each I_r . Equivalently, each probability mass $m_r(\theta)$ originally concentrated at $\hat{v}_r$ is redistributed over I_r in the same proportion as the prior F . Let $\tilde{\theta}_\delta(\theta)$ be the index of this lifted posterior, i.e., the unique solution to

$$\int_0^1 ((v - \tilde{\theta}_\delta(\theta))_+ - c) d\tilde{G}_{v|\theta}(v) = 0 .$$

Thus, each row θ of G_δ^* is transformed as follows: mass at each midpoint $\widehat{v}_r$ is spread back to I_r , and the whole row is then moved vertically from θ to the recomputed index $\widetilde{\theta}_\delta(\theta)$. Formally, define

$$G(A \times B) \triangleq \int_{-c}^{1-c} \mathbf{1}\{\widetilde{\theta}_\delta(\theta) \in B\} \widetilde{G}_{v|\theta}(A) \, dG_{\delta,\theta}^*(\theta)$$

for all measurable $A \subseteq [0, 1]$ and $B \subseteq [-c, 1 - c]$. It can be easily verified that $G \in \mathcal{G}(F, c)$.

We next show that the vertical and horizontal movements are both small. For every $v \in I_r$, we have $|v - \widehat{v}_r| \leq \delta$. Since G_δ^* is feasible, for $G_{\delta,\theta}^*$ -almost every row θ , $\sum_{r=1}^M m_r(\theta) ((\widehat{v}_r - \theta)_+ - c) = 0$. For every $v \in I_r$, we have

$$(v - (\theta + \delta))_+ \leq (\widehat{v}_r - \theta)_+, \quad (v - (\theta - \delta))_+ \geq (\widehat{v}_r - \theta)_+.$$

Integrating with respect to the lifted posterior $\widetilde{G}_{v|\theta}$ gives

$$\int_0^1 (v - (\theta + \delta))_+ \, d\widetilde{G}_{v|\theta}(v) \leq c, \quad \int_0^1 (v - (\theta - \delta))_+ \, d\widetilde{G}_{v|\theta}(v) \geq c.$$

By monotonicity of the index equation in the index variable,

$$|\widetilde{\theta}_\delta(\theta) - \theta| \leq \delta \quad \text{for } G_{\delta,\theta}^*\text{-almost every } \theta.$$

It remains to compare payoffs. For every $v \in I_r$ and $G_{\delta,\theta}^*$ -almost every θ , we have

$$\left| \min\{v, \widetilde{\theta}_\delta(\theta)\} - \min\{\widehat{v}_r, \theta\} \right| \leq \delta.$$

Since the interim utility function $u(\kappa)$ is L -Lipschitz continuous, we can bound the utility loss

$$u\left(\min\{v, \widetilde{\theta}_\delta(\theta)\}\right) \geq u\left(\min\{\widehat{v}_r, \theta\}\right) - L\delta.$$

Applying this pointwise payoff bound on each interval I_r , we obtain

$$\begin{aligned} U(G) &= \int_{-c}^{1-c} \sum_{r=1}^M m_r(\theta) \int_{I_r} u\left(\min\{v, \widetilde{\theta}_\delta(\theta)\}\right) \frac{dF(v)}{F(I_r)} \, dG_{\delta,\theta}^*(\theta) \\ &\geq \int_{-c}^{1-c} \sum_{r=1}^M m_r(\theta) [u\left(\min\{\widehat{v}_r, \theta\}\right) - L\delta] \, dG_{\delta,\theta}^*(\theta) \\ &= U(G_\delta^*) - L\delta = \text{OPT}_\delta - L\delta. \end{aligned}$$

Since G^* is optimal for the original problem, we conclude that

$$\text{OPT} \geq U(G) \geq \text{OPT}_\delta - L\delta.$$

Combining the two inequalities from steps 1.1 and 1.2 yields

$$|\text{OPT}_\delta - \text{OPT}| \leq L\delta.$$

Step 2: Discretizing the index support. Consider the discrete prior F_δ constructed in Step 1. Fix an index-discretization parameter $\epsilon \in (0, 1)$. Recall that the discretized prior constructed in the previous step is $F_\delta = \sum_{i=1}^M f_i^\delta \cdot \delta_{\widehat{v}_i}$ and $f_i^\delta \triangleq F(I_i)$. For each grid value $\widehat{v}_i$ and each index $\theta \in [-c, 1 - c]$, define these two functions for convenience $q_i(\theta) \triangleq (\widehat{v}_i - \theta)_+ - c$ and $p_i(\theta) \triangleq u(\min\{\widehat{v}_i, \theta\})$. Let $\Theta_\delta \triangleq [\widehat{v}_1 - c, \widehat{v}_M - c]$ be the effective index domain under the discretized prior.

Step 2.1: Constructing the index grid. We first discretize the vertical axis of the interim utility function $u(\kappa)$ at granularity ϵ . Let $\ell_k \triangleq k\epsilon$ for $k = 0, 1, \dots, \lceil 1/\epsilon \rceil$. For every level $\ell_k \leq u(\widehat{v}_M - c)$, define the generalized inverse $\xi_k \triangleq \inf\{\theta \in \Theta_\delta : u(\theta) \geq \ell_k\}$. Since the interim utility $u(\kappa)$ is continuous and nondecreasing, we have $u(\xi_k) \geq \ell_k$ for each k .

Then, the full index grid is constructed as

$$\Gamma \triangleq \{\widehat{v}_i - c : i \in [M]\} \cup (\{\widehat{v}_i : i \in [M]\} \cap \Theta_\delta) \cup \{\xi_k : k = 0, 1, \dots, \lceil 1/\epsilon \rceil\}.$$

After sorting and removing duplicates, write

$$\Gamma = \{\theta_1, \dots, \theta_R\}, \quad \widehat{v}_1 - c = \theta_1 < \dots < \theta_R = \widehat{v}_M - c.$$

The grid size satisfies $|\Gamma| \triangleq R \leq 2M + \lceil 1/\epsilon \rceil + 3$, which is finite.

We illustrate why constructing such an index grid. The points $\widehat{v}_i$ are added so that the map $\theta \mapsto \min\{\widehat{v}_i, \theta\}$ has no kink inside any grid interval. The points $\widehat{v}_i - c$ are added so that the coefficient $q_i(\theta) = (\widehat{v}_i - \theta)_+ - c$ does not change sign inside any open interval between two consecutive grid points. The points ξ_k ensure the one-sided payoff approximation used in the rounding step. Specifically, if $\alpha < \beta$ are two consecutive points in Γ , then for every $\theta \in [\alpha, \beta)$,

$$u(\theta) \leq u(\alpha) + \epsilon.$$

To see this, if $u(\theta) > u(\alpha) + \epsilon$ for some $\theta \in [\alpha, \beta)$, then there is a level ℓ_k such that $u(\alpha) < \ell_k \leq u(\theta)$. By definition of ξ_k , we have $\xi_k \leq \theta < \beta$. If $\xi_k \leq \alpha$, then monotonicity implies $u(\alpha) \geq u(\xi_k) \geq \ell_k$, a contradiction. Hence $\xi_k \in (\alpha, \beta)$, contradicting the fact that α and β are consecutive grid points.

Step 2.2: Defining the finite LP. We now restrict the index support to the finite grid Γ . Let $\mathbf{x} = (x_{ir})_{i \in [M], r \in [R]}$ be the decision variable, where $x_{ir} \geq 0$ denotes the probability mass assigned to value $\widehat{v}_i$ and index θ_r . The discretized problem is the following finite-dimensional LP:

$$\begin{aligned} \max_{\mathbf{x} \geq 0} \quad & \sum_{i=1}^M \sum_{r=1}^R u(\min\{\widehat{v}_i, \theta_r\}) x_{ir} \\ \text{s.t.} \quad & \sum_{r=1}^R x_{ir} = f_i^\delta, \quad \forall i \in [M], \\ & \sum_{i=1}^M ((\widehat{v}_i - \theta_r)_+ - c) x_{ir} = 0, \quad \forall r \in [R]. \end{aligned} \tag{P_{BR}^{\delta, \epsilon}}$$

The LP is feasible since the full-revelation strategy belongs to its feasible region: for each value $\widehat{v}_i$, assign mass f_i^δ to the index $\widehat{v}_i - c$. Let $G_{\delta, \epsilon}^*$ be an optimal solution to $(\mathcal{P}_{BR}^{\delta, \epsilon})$, and let $\text{OPT}_{\delta, \epsilon} \triangleq U(G_{\delta, \epsilon}^*)$ be its optimal value. Every feasible solution of $(\mathcal{P}_{BR}^{\delta, \epsilon})$ is feasible for $(\mathcal{P}_{BR}^\delta)$, because it is simply a feasible joint distribution under F_δ with finite index support. Therefore,

$$\text{OPT}_{\delta, \epsilon} \leq \text{OPT}_\delta.$$

Step 2.3: Rounding an optimal solution to the index grid. Let $G_\delta^* \in \mathcal{G}(F_\delta, c)$ be an optimal solution to the problem $(\mathcal{P}_{BR}^\delta)$ under discrete prior F_δ . In this step, we round all mass of G_δ^* on non-grid index rows to the finite grid Γ , while preserving feasibility.

For each grid value $\widehat{v}_i$, let γ_i^* be the finite Borel measure on Θ_δ induced by G_δ^* :

$$\gamma_i^*(A) \triangleq G_\delta^*(\{\widehat{v}_i\} \times A), \quad \forall A \in \mathcal{B}(\Theta_\delta).$$

It suffices to describe the rounding operation on one open interval between two consecutive grid points. Fix $\alpha < \beta$ in Γ with no grid point between them, and let $J \triangleq (\alpha, \beta)$. We will move all mass with indices in J to the two endpoint rows α and β .

Let $M_J \triangleq \sum_{i=1}^M \gamma_i^*|_J$ be the total index-mass measure on J . If $M_J(J) = 0$, there is nothing to move. Otherwise, for each $i \in [M]$, define the row composition at θ

$$z_i(\theta) \triangleq \frac{d\gamma_i^*|_J}{dM_J}(\theta) .$$

For M_J -almost every $\theta \in J$, the vector $z(\theta) = (z_1(\theta), \dots, z_M(\theta))$ satisfies

$$\sum_{i=1}^M z_i(\theta) = 1 , \quad \sum_{i=1}^M q_i(\theta) z_i(\theta) = 0 .$$

The key observation is that each such feasible row can be split into two feasible endpoint rows at α and β . Since Γ contains all kink points $\widehat{v}_i$ and all zero-crossing points $\widehat{v}_i - c$, every q_i is affine on $[\alpha, \beta]$ and has a fixed sign on J . Define

$$P \triangleq \{i \in [M] : q_i(\theta) > 0\} , \quad N \triangleq \{j \in [M] : q_j(\theta) < 0\} .$$

Then these two sets, P and N , form a partition $[M]$, and row feasibility (the index equation in Definition 1) implies that the positive and negative contributions balance:

$$A(\theta) \triangleq \sum_{i \in P} q_i(\theta) z_i(\theta) = \sum_{j \in N} (-q_j(\theta)) z_j(\theta) > 0 .$$

For each pair $(i, j) \in P \times N$, define

$$a_{ij}(\theta) \triangleq \frac{z_i(\theta) z_j(\theta)}{A(\theta)} .$$

Also define the two-value balanced package $r^{ij}(\theta) \in \mathbb{R}_+^M$ by

$$r_k^{ij}(\theta) \triangleq \begin{cases} -q_j(\theta) & \text{if } k = i , \\ q_i(\theta) & \text{if } k = j , \\ 0 & \text{otherwise .} \end{cases}$$

This package is feasible at row θ , because $\sum_k q_k(\theta) r_k^{ij}(\theta) = 0$. Moreover, the row vector $z(\theta)$ can be decomposed as

$$z(\theta) = \sum_{i \in P, j \in N} a_{ij}(\theta) r^{ij}(\theta) .$$

To see this, the i -th coordinate for $i \in P$ equals

$$\sum_{j \in N} a_{ij}(\theta) (-q_j(\theta)) = \frac{z_i(\theta)}{A(\theta)} \sum_{j \in N} (-q_j(\theta)) z_j(\theta) = z_i(\theta) ,$$

and the verification for $j \in N$ is identical.

We now move each balanced package to the endpoints. Write

$$\theta = \omega(\theta)\alpha + (1 - \omega(\theta))\beta , \quad \omega(\theta) \triangleq \frac{\beta - \theta}{\beta - \alpha} .$$

Since each q_i is affine on $[\alpha, \beta]$, we have

$$r^{ij}(\theta) = \omega(\theta) r^{ij}(\alpha) + (1 - \omega(\theta)) r^{ij}(\beta) .$$

Therefore, define

$$z^\alpha(\theta) \triangleq \omega(\theta) \sum_{i \in P, j \in N} a_{ij}(\theta) r^{ij}(\alpha) , \quad z^\beta(\theta) \triangleq (1 - \omega(\theta)) \sum_{i \in P, j \in N} a_{ij}(\theta) r^{ij}(\beta) .$$

Then $z(\theta) = z^\alpha(\theta) + z^\beta(\theta)$. Moreover, $z^\alpha(\theta)$ is feasible at row α , and $z^\beta(\theta)$ is feasible at row β , because each package $r^{ij}(\alpha)$ and $r^{ij}(\beta)$ is feasible at the corresponding endpoint row.

We now apply this local splitting operation to every open interval between consecutive grid points. Let $J_r \triangleq (\theta_r, \theta_{r+1})$ for $r \in [R-1]$. For each J_r , let M_r be the total index-mass measure on J_r , and let $z^{r,L}$ and $z^{r,R}$ be the endpoint split vectors sent to θ_r and θ_{r+1} , respectively. Define the rounded measures $\hat{\gamma}_i$ by

$$\hat{\gamma}_i \triangleq \gamma_i^*|_\Gamma + \sum_{r=1}^{R-1} \left(\int_{J_r} z_i^{r,L}(\theta) \, dM_r(\theta) \right) \delta_{\theta_r} + \sum_{r=1}^{R-1} \left(\int_{J_r} z_i^{r,R}(\theta) \, dM_r(\theta) \right) \delta_{\theta_{r+1}} .$$

Finally, define the rounded joint distribution $\hat{G}_{\delta,\epsilon}$ by

$$\hat{G}_{\delta,\epsilon}(\{\hat{v}_i\} \times A) \triangleq \hat{\gamma}_i(A) , \quad \forall A \in \mathcal{B}(\Theta_\delta) .$$

By construction, $\hat{G}_{\delta,\epsilon}$ is supported on the grid points $\{\hat{v}_1, \dots, \hat{v}_M\} \times \Gamma$.

We verify feasibility. First, the value marginal is preserved. For every $i \in [M]$,

$$\begin{aligned} \hat{\gamma}_i(\Theta_\delta) &= \gamma_i^*(\Gamma) + \sum_{r=1}^{R-1} \int_{J_r} \left(z_i^{r,L}(\theta) + z_i^{r,R}(\theta) \right) \, dM_r(\theta) \\ &= \gamma_i^*(\Gamma) + \sum_{r=1}^{R-1} \int_{J_r} z_i(\theta) \, dM_r(\theta) = \gamma_i^*(\Theta_\delta) = f_i^\delta . \end{aligned}$$

Second, the row constraints are preserved. Since $\hat{G}_{\delta,\epsilon}$ is supported only on Γ , it suffices to check the constraint at each grid row θ_r . The mass originally placed on θ_r is feasible because G_δ^* is feasible. Every piece of mass moved into θ_r from a neighboring interval is feasible at θ_r by the endpoint-splitting construction. Hence, writing $\hat{x}_{ir} \triangleq \hat{\gamma}_i(\{\theta_r\})$, we have

$$\sum_{i=1}^M ((\hat{v}_i - \theta_r)_+ - c) \hat{x}_{ir} = 0 , \quad \forall r \in [R] .$$

Therefore, $\hat{G}_{\delta,\epsilon}$ is feasible for the finite LP $(\mathcal{P}_{\text{BR}}^{\delta,\epsilon})$.

Step 2.4: Bounding the rounding loss. We now bound the payoff loss caused by the rounding operation for every probability mass. Fix a grid interval (α, β) and a nonempty row $\theta \in (\alpha, \beta)$ with row vector $z(\theta)$. This row is replaced by a value vector $z^\alpha(\theta)$ at row α and a value vector $z^\beta(\theta)$ at row β . By construction of the grid, for value $\hat{v}_i$,

$$p_i(\alpha) \geq p_i(\theta) - \epsilon , \quad p_i(\beta) \geq p_i(\theta) .$$

Indeed, the payoff grid ensures that moving to the left endpoint loses at most ϵ . Therefore, the payoff after rounding this row is at least

$$\begin{aligned} \sum_{i=1}^M p_i(\alpha) z_i^\alpha(\theta) + \sum_{i=1}^M p_i(\beta) z_i^\beta(\theta) &\geq \sum_{i=1}^M (p_i(\theta) - \epsilon) z_i^\alpha(\theta) + \sum_{i=1}^M p_i(\theta) z_i^\beta(\theta) \\ &= \sum_{i=1}^M p_i(\theta) \left(z_i^\alpha(\theta) + z_i^\beta(\theta) \right) - \epsilon \sum_{i=1}^M z_i^\alpha(\theta) \\ &\geq \sum_{i=1}^M p_i(\theta) z_i(\theta) - \epsilon \sum_{i=1}^M z_i(\theta) . \end{aligned}$$

Thus the loss on this row is at most $\epsilon \sum_{i=1}^M z_i(\theta)$. Summing over all rows and all grid intervals, and using that the total probability mass is one, the aggregate payoff loss is at most $\epsilon = \epsilon$. Hence

$$\text{OPT}_{\delta, \epsilon} \geq U(\hat{G}_{\delta, \epsilon}) \geq U(G_{\delta}^*) - \epsilon = \text{OPT}_{\delta} - \epsilon .$$

Combining this with $\text{OPT}_{\delta, \epsilon} \leq \text{OPT}_{\delta}$, we conclude

$$\text{OPT}_{\delta} - \epsilon \leq \text{OPT}_{\delta, \epsilon} \leq \text{OPT}_{\delta} .$$

Step 3: Lifting the finite-LP solution back to the original prior. We now construct a feasible solution under the original prior F from the finite-LP solution $G_{\delta, \epsilon}^*$. Let $\mathbf{x}^* = (x_{ir}^*)_{i \in [M], r \in [R]}$ be the optimal solution to the finite LP $(\mathcal{P}_{\text{BR}}^{\delta, \epsilon})$, so that $G_{\delta, \epsilon}^*$ assigns mass x_{ir}^* to the value-index pair $(\hat{v}_i, \theta_r)$ for each $i \in [M]$ and $r \in [R]$.

For each grid row θ_r , write $Y_r \triangleq \sum_{i=1}^M x_{ir}^*$. If $Y_r = 0$, this row carries no mass and can be ignored. If $Y_r > 0$, we lift the posterior on row θ_r by spreading the mass at each midpoint $\hat{v}_i$ back to the interval I_i according to the conditional distribution of F on I_i . Specifically, define a posterior distribution $\tilde{G}_{v|r}$ by, for every measurable set $A \subseteq [0, 1]$,

$$\tilde{G}_{v|r}(A) \triangleq \sum_{i=1}^M \frac{x_{ir}^*}{Y_r} \frac{F(A \cap I_i)}{F(I_i)} .$$

Since F is strictly increasing, $F(I_i) > 0$ for every $i \in [M]$. Let $\tilde{\theta}_r$ be the index of this lifted posterior, i.e., the unique solution to the index equation

$$\int_0^1 \left((v - \tilde{\theta}_r)_+ - c \right) d\tilde{G}_{v|r}(v) = 0 .$$

Thus, each grid row θ_r is transformed as follows: every mass x_{ir}^* originally placed at the value midpoint $\hat{v}_i$ is spread back over I_i , and the entire row is moved vertically from the original index θ_r to the newly recomputed index $\tilde{\theta}_r$. Formally, define $\tilde{G}$ by

$$\tilde{G}(A \times B) \triangleq \sum_{r: Y_r > 0} \mathbf{1}\{\tilde{\theta}_r \in B\} \sum_{i=1}^M x_{ir}^* \frac{F(A \cap I_i)}{F(I_i)}$$

for all measurable $A \subseteq [0, 1]$ and $B \subseteq [-c, 1 - c]$.

We first verify that $\tilde{G} \in \mathcal{G}(F, c)$. For the value marginal, for every measurable set $A \subseteq [0, 1]$,

$$\tilde{G}(A \times [-c, 1 - c]) = \sum_{r=1}^R \sum_{i=1}^M x_{ir}^* \frac{F(A \cap I_i)}{F(I_i)} = \sum_{i=1}^M \frac{F(A \cap I_i)}{F(I_i)} \sum_{r=1}^R x_{ir}^* = \sum_{i=1}^M F(A \cap I_i) = F(A) .$$

Hence the value marginal of $\tilde{G}$ is F . The row-wise index constraint also holds by construction. Indeed, for any bounded measurable function $\varphi : [-c, 1 - c] \rightarrow \mathbb{R}$,

$$\int_{[0, 1] \times [-c, 1 - c]} \varphi(\epsilon) \left((v - \epsilon)_+ - c \right) d\tilde{G}(v, \epsilon) = \sum_{r: Y_r > 0} \varphi(\tilde{\theta}_r) Y_r \int_0^1 \left((v - \tilde{\theta}_r)_+ - c \right) d\tilde{G}_{v|r}(v) = 0 .$$

Therefore, the constructed strategy $\tilde{G}$ is feasible for the original problem $(\mathcal{P}_{\text{BR}})$.

We next show that the vertical and horizontal movements are both small. For every $v \in I_i$, we have $|v - \hat{v}_i| \leq \delta$. Since $G_{\delta, \epsilon}^*$ is feasible for the finite LP, for every grid row θ_r with $Y_r > 0$,

$$\sum_{i=1}^M \frac{x_{ir}^*}{Y_r} (\hat{v}_i - \theta_r)_+ = c .$$

For every $v \in I_i$, we have

$$(v - (\theta_r + \delta))_+ \leq (\widehat{v}_i - \theta_r)_+ , \quad (v - (\theta_r - \delta))_+ \geq (\widehat{v}_i - \theta_r)_+ .$$

Integrating with respect to the lifted posterior $\widetilde{G}_{v|r}$ gives

$$\int_0^1 (v - (\theta_r + \delta))_+ d\widetilde{G}_{v|r}(v) \leq c , \quad \int_0^1 (v - (\theta_r - \delta))_+ d\widetilde{G}_{v|r}(v) \geq c .$$

By monotonicity of the index equation in the index variable,

$$|\widetilde{\theta}_r - \theta_r| \leq \delta \quad \text{for every } r \text{ with } Y_r > 0 .$$

It remains to compare payoffs. For every $v \in I_i$ and every row r with $Y_r > 0$,

$$\left| \min\{v, \widetilde{\theta}_r\} - \min\{\widehat{v}_i, \theta_r\} \right| \leq \delta .$$

Since the interim utility function $u(\kappa)$ is L -Lipschitz continuous,

$$u(\min\{v, \widetilde{\theta}_r\}) \geq u(\min\{\widehat{v}_i, \theta_r\}) - L\delta .$$

Using the definition of $\widetilde{G}$, we obtain the following utility bound

$$\begin{aligned} U(\widetilde{G}) &= \sum_{r: Y_r > 0} \sum_{i=1}^M x_{ir}^* \int_{I_i} u(\min\{v, \widetilde{\theta}_r\}) \frac{dF(v)}{F(I_i)} \\ &\geq \sum_{r: Y_r > 0} \sum_{i=1}^M x_{ir}^* [u(\min\{\widehat{v}_i, \theta_r\}) - L\delta] = U(G_{\delta, \epsilon}^*) - L\delta = \text{OPT}_{\delta, \epsilon} - L\delta . \end{aligned}$$

Combining this inequality with the bounds from Steps 1 and 2 gives

$$U(\widetilde{G}) \geq \text{OPT} - \epsilon - 2L\delta .$$

Thus, the finite-LP solution $G_{\delta, \epsilon}^*$ can be lifted back to the original continuous-prior problem ($\mathcal{P}_{\text{BR}}$) with only an additional loss of $L\delta$.

To obtain a final additive error of at most $\bar{\epsilon} > 0$, choose the index-discretization parameter and the value-grid size so that $\epsilon = \bar{\epsilon}/2$ and $\delta \leq \bar{\epsilon}/4L$. With these choices, we have $\epsilon + 2L\delta \leq \bar{\epsilon}$, and hence the constructed strategy $\widetilde{G}$ in above steps satisfies $U(\widetilde{G}) \geq \text{OPT} - \bar{\epsilon}$.

The number of value grid points is $M = O(1/\delta)$, and the number of index grid points satisfies $R \leq 2M + \lceil 1/\epsilon \rceil + 3$. Therefore, the finite LP has MR variables and $M + R$ equality constraints, both polynomial in L and $1/\bar{\epsilon}$. Assuming that the interval masses $F(I_i)$ can be computed in polynomial time, the whole procedure runs in polynomial time.

Therefore, for every target accuracy $\bar{\epsilon} > 0$, the procedure outputs in polynomial time a feasible solution $\widetilde{G} \in \mathcal{G}(F, c)$ such that $U(\widetilde{G}) \geq \text{OPT} - \bar{\epsilon}$. This proves that the sender's persuasion problem admits a fully polynomial additive approximation scheme. $\square$

D Equilibrium Existence of the Competitive Information Design Game

In this section, we establish the existence of an equilibrium in the game among the senders, under some mild assumptions.

Theorem 5 (Equilibrium Existence). *There exists an equilibrium in the game among the senders if each sender i 's prior distribution F_i is atomless over $[0, 1]$.*

We proceed with the proof in the following three steps. Briefly speaking, in Step 1, we construct a finite game called m -discrete approximation game for each granularity $m \in \mathbb{Z}^+$, and directly apply Nash's Theorem to show the existence of a discrete equilibrium. In Step 2, we show that the sequence of discrete equilibria has a subsequence that weakly converges to some distribution profile, where any of the distributions forms a feasible action for each sender. In Step 3, we prove that the limit profile indeed constitutes an equilibrium of the original game through establishing the convergence of utilities along with the convergence of equilibrium strategies. For ease of presentation, we assume all senders share a common inspection cost c , that is $c_1 = \dots = c_N = c$.

Step 1: Constructing m -discrete approximation games. By Lemma 4, each sender i 's strategy space $\mathcal{G}(F_i, c)$ is a compact and convex set containing all feasible 2-D distributions. By Krein-Milman Theorem, any convex and compact set is the closed convex hull of its extreme points; moreover, each element within the set can be represented as a convex combination of its extreme points. Thus, each sender's strategy space can be precisely characterized by its extreme points. Recall from Definition 4 that the space $\mathcal{G}(F_i, c)$ is subject to an infinite number of constraints; the convex space $\mathcal{G}(F_i, c)$ may therefore have an infinite number of extreme points. Therefore, even if one takes the set of these extreme points as the action set, Nash's theorem does not apply. To address this problem, we construct finite games that not only discretize the support of the value and the index, but also modify constraints in Definition 4. Each finite game is parameterized by an integer $m \in \mathbb{Z}^+$.

Definition 6 (m -Discrete Approximation Game). Fix any $m \in \mathbb{Z}^+$, let $V^m \triangleq \{0, \frac{1}{2^m}, \dots, 1\}$ and $\Theta^m \triangleq \{-c, \frac{1}{2^m} - c, \dots, 1 - c\}$ be the discretized support for the value and the index, respectively. The m -discrete approximation game is as follows:

- **Strategy space:** For each sender i , the strategy space $\mathcal{S}_i^m$ is a subset of distributions on $V^m \times \Theta^m$. Each $\mathbf{p} \in \mathcal{S}_i^m$, with $p_{i,j}$ denoting the probability on the point $(\frac{i}{2^m}, \frac{j}{2^m} - c)$, satisfies

$$\sum_{i=0}^{2^m} p_{i,j} \cdot \left[\frac{i}{2^m} - \frac{j}{2^m} + c \right]_+ = c \cdot \sum_{i=0}^{2^m} p_{i,j}, \quad \forall j \in \{0\} \cup [2^m]; \quad (21)$$

$$\sum_{j=0}^{2^m} p_{i,j} = F_i \left(\frac{i}{2^m} \right) - F_i \left(\left[\frac{i-1}{2^m} \right]_+ \right), \quad \forall i \in \{0\} \cup [2^m]. \quad (22)$$

- **Action space:** Each sender i 's action space, denoted by $\mathcal{A}_i^m$, is defined as the set of all extreme points of $\mathcal{S}_i^m$.
- **Utility:** Given a strategy profile $(G_1^m, \dots, G_N^m)$ where each $G_i^m \in \mathcal{A}_i^m$, the receiver receives $(v_i^m, \theta_i^m) \sim G_i^m$ for each i , and chooses a sender whose amortized value $\kappa_i^m \triangleq \min(v_i^m, \theta_i^m)$ is non-negative and the largest. If all amortized values are negative, the receiver chooses no one. If there is a tie among multiple senders, break the tie uniformly at random.¹⁰ A sender gets utility 1 if chosen by the receiver, and 0 otherwise.

Lemma 18. For any $m \in \mathbb{Z}^+$, each sender i 's strategy space $\mathcal{S}_i^m$ is non-empty, convex, and compact. Each sender i 's action space $\mathcal{A}_i^m$ is finite.

Proof of Lemma 18. First, we prove that the space $\mathcal{S}_i^m$ is non-empty. For any $m \in \mathbb{Z}^+$ and each sender i , the full-revelation strategy belongs to the space $\mathcal{S}_i^m$, which makes the space $\mathcal{S}_i^m$ non-empty.

Second, we prove the space $\mathcal{S}_i^m$ is convex. For any $m \in \mathbb{Z}^+$, consider any pair of feasible strategies $G^m, H^m \in \mathcal{S}_i^m$ (with probability mass $g_{i,j}$ and $h_{i,j}$ on the point $(\frac{i}{2^m}, \frac{j}{2^m} - c)$)

¹⁰ Here we adopt the uniformly random tie-breaking rule for convenience of presentation. As we remarked above, the proof goes through for a host of other tie-breaking rules.

respectively) and any $\lambda \in [0, 1]$, we construct the convex combination T^m where each mass $t_{i,j} = \lambda \cdot g_{i,j} + (1 - \lambda) \cdot h_{i,j}$. It is obvious that for each $j \in \{0\} \cup [2^m]$,

$$\begin{aligned} & \sum_{i=0}^{2^m} t_{i,j} \cdot \max \left\{ \frac{i}{2^m} - \frac{j}{2^m} + c, 0 \right\} \\ &= \lambda \cdot \sum_{i=0}^{2^m} g_{i,j} \cdot \max \left\{ \frac{i}{2^m} - \frac{j}{2^m} + c, 0 \right\} + (1 - \lambda) \cdot \sum_{i=0}^{2^m} h_{i,j} \cdot \max \left\{ \frac{i}{2^m} - \frac{j}{2^m} + c, 0 \right\} \\ &= \lambda \cdot c \cdot \sum_{i=0}^{2^m} g_{i,j} + (1 - \lambda) \cdot c \cdot \sum_{i=0}^{2^m} h_{i,j} = c \cdot \sum_{i=0}^{2^m} t_{i,j}, \end{aligned}$$

and for each $i \in \{0\} \cup [2^m]$,

$$\sum_{j=0}^{2^m} t_{i,j} = \lambda \cdot \sum_{j=0}^{2^m} g_{i,j} + (1 - \lambda) \cdot \sum_{j=0}^{2^m} h_{i,j} = F\left(\frac{i}{2^m}\right) - F\left(\max\left\{0, \frac{i-1}{2^m}\right\}\right).$$

So we know that the convex combination T^m also belongs to the space $\mathcal{S}_i^m$, which makes the space $\mathcal{S}_i^m$ convex.

Third, we prove that the space $\mathcal{S}_i^m$ is compact. Since the space $\mathcal{S}_i^m$ is indeed a measure space over $V^m \times \Theta^m$, the space $\mathcal{S}_i^m$ is bounded. For any $m \in \mathbb{Z}^+$ and any sender i , we assume that sequence $\{G^k\}_{k \in \mathbb{Z}^+}$ converges to some discrete distribution G where each $G^k \in \mathcal{S}_i^m$. So the sequence $\{p_{i,j}^k\}_{k \in \mathbb{Z}^+}$ converges to $p_{i,j}$ where $p_{i,j}^k$ denotes the probability measure of strategy G^k on the point $(\frac{i}{2^m}, \frac{j}{2^m} - c)$ for any $i, j \in \{0\} \cup [2^m]$, and $p_{i,j}$ denotes the probability measure of strategy G on the same point. This implies that there exists $K > 0$ such that for any $k > K$ and any $\epsilon > 0$, we have $|p_{i,j}^k - p_{i,j}| < \epsilon$ for any $i, j \in \{0\} \cup [2^m]$. If we assume that there exist some m and $j \in \{0\} \cup [2^m]$ such that $\sum_{i=0}^{2^m} p_{i,j} \cdot \max\left\{\frac{i}{2^m} - \frac{j}{2^m} + c, 0\right\} = d \cdot \sum_{i=0}^{2^m} p_{i,j} \neq c \cdot \sum_{i=0}^{2^m} p_{i,j}$. There exists $K_1 > 0$ such that for any $k > K_1$, we have

$$\left| \sum_{i=0}^{2^m} p_{i,j}^k \cdot \max\left\{\frac{i}{2^m} - \frac{j}{2^m} + c, 0\right\} - \sum_{i=0}^{2^m} p_{i,j} \cdot \max\left\{\frac{i}{2^m} - \frac{j}{2^m} + c, 0\right\} \right| < \frac{1}{3} \left| d \cdot \sum_{i=0}^{2^m} p_{i,j} - c \cdot \sum_{i=0}^{2^m} p_{i,j} \right|.$$

There exists $K_2 > 0$ such that for any $k > K_2$, we have

$$\left| c \cdot \sum_{i=0}^{2^m} p_{i,j}^k - c \cdot \sum_{i=0}^{2^m} p_{i,j} \right| < \frac{1}{3} \left| d \cdot \sum_{i=0}^{2^m} p_{i,j} - c \cdot \sum_{i=0}^{2^m} p_{i,j} \right|.$$

Thus, we know for any $k > \max\{K_1, K_2\}$, we have

$$\left| \sum_{i=0}^{2^m} p_{i,j}^k \cdot \max\left\{\frac{i}{2^m} - \frac{j}{2^m} + c, 0\right\} - c \cdot \sum_{i=0}^{2^m} p_{i,j}^k \right| > \frac{1}{3} \left| d \cdot \sum_{i=0}^{2^m} p_{i,j} - c \cdot \sum_{i=0}^{2^m} p_{i,j} \right|,$$

which forms a contradiction with the fact that $G^k \in \mathcal{S}_i^m$. Furthermore, if we assume that there exist some m and $j \in \{0\} \cup [2^m]$ such that $\sum_{j=0}^{2^m} p_{i,j} \neq F\left(\frac{i}{2^m}\right) - F\left(\max\left\{0, \frac{i-1}{2^m}\right\}\right)$, then we can achieve a contradiction in the same manner. So we have proved that the space $\mathcal{S}_i^m$ is closed. Then we can directly make space $\mathcal{S}_i^m$ compact through the Heine-Borel Theorem.

Last, we prove that space $\mathcal{A}_i^m$ is finite. By Definition 6, set $\mathcal{S}_i^m$ is formed by $O(2^m)$ linear constraints, which implies that there are a finite number of extreme points of set $\mathcal{S}_i^m$. Thus, the action space $\mathcal{A}_i^m$ is finite. $\square$

Step 2: Showing the limit of equilibrium strategies is feasible. By Lemma 18, each m -discrete approximation game is finite, there being a finite number of senders, and each sender i having a finite action space. Thus, Nash's Theorem applies, and there is an equilibrium $(\tilde{G}_1^m, \dots, \tilde{G}_N^m)$, where for each sender i , $\tilde{G}_i^m \in \mathcal{S}_i^m$ is a mixed equilibrium strategy. For each sender i , these equilibrium strategies form a sequence of 2-D distributions: $\{\tilde{G}_i^m\}_{m \in \mathbb{Z}^+}$. We want to establish the convergence of this sequence, so we first need to introduce the following Helly's Selection Theorem.

Lemma 19 (Helly's Selection Theorem). *Let $\{G^m\}_{m \in \mathbb{Z}^+}$ be a sequence of CDFs which is tight,¹¹ then there exists a subsequence $\{m(k)\}_{k \in \mathbb{Z}^+} \subseteq \mathbb{Z}^+$ such that $\{G^{m(k)}\}_{k \in \mathbb{Z}^+}$ weakly converges to a certain CDF G .¹²*

By Helly's Selection Theorem, there is a subsequence that weakly converges to a 2-D distribution; we further show that this limit is a feasible for each sender i .

Lemma 20. *There exists a subsequence $\{m(k)\}_{k \in \mathbb{Z}^+} \subseteq \mathbb{Z}^+$ such that the sequence $\{(\tilde{G}_i^{m(k)})\}_{k \in \mathbb{Z}^+}$ weakly converges to a certain 2-D distribution $\tilde{G}_i$. Furthermore, $\tilde{G}_i$ is a feasible strategy for sender i in the original game.*

Proof of Lemma 20. For any $m \in \mathbb{Z}^+$, space $\mathcal{S}_i^m$ is a measure space over $[0, 1] \times [-c, 1 - c]$, which makes space $\mathcal{S}_i^m$ tight. By Lemma 19, for each sender $i \in [N]$, there exists a subsequence $\{m_i(k)\}_{k \in \mathbb{Z}^+} \subseteq \mathbb{Z}^+$ such that $\{\tilde{G}_i^{m_i(k)}\}_{k \in \mathbb{Z}^+}$ weakly converges to a certain 2-d CDF $\tilde{G}_i$. This means that, for any continuous point $(\hat{v}, \hat{\theta})$ of distribution $\tilde{G}_i$, the sequence $\{\tilde{G}_i^{m_i(k)}(\hat{v}, \hat{\theta})\}_{k \in \mathbb{Z}^+}$ converges to $\tilde{G}_i(\hat{v}, \hat{\theta})$. Next, it can be shown that, there exists a common subsequence $\{m(k)\}_{k \in \mathbb{Z}^+}$ such that for each sender i , the sequence $\{\tilde{G}_i^{m(k)}\}_{k \in \mathbb{Z}^+}$ weakly converges to a certain CDF $\tilde{G}_i$. To see this, we first find a subsequence $\{m_1(k)\}_{k \in \mathbb{Z}^+}$ of $\mathbb{Z}^+$ such that $\{\tilde{G}_1^{m_1(k)}\}_{k \in \mathbb{Z}^+}$ weakly converges to $\tilde{G}_1$. We then find a subsequence $\{m_2(k)\}_{k \in \mathbb{Z}^+}$ of $\{m_1(k)\}_{k \in \mathbb{Z}^+}$ such that $\{\tilde{G}_2^{m_2(k)}\}_{k \in \mathbb{Z}^+}$ weakly converges to $\tilde{G}_2$. It is known that if a sequence converges, any subsequence of it also converges to the same limit. Hence, $\{\tilde{G}_1^{m_2(k)}\}_{k \in \mathbb{Z}^+}$ also weakly converges to $\tilde{G}_1$. In the same manner, we obtain a common subsequence $\{m(k)\}_{k \in \mathbb{Z}^+}$ of $\mathbb{Z}^+$ such that for each sender i , the sequence $\{\tilde{G}_i^{m(k)}\}_{k \in \mathbb{Z}^+}$ weakly converges to a certain CDF $\tilde{G}_i$.

Then, we are going to show that the limit 2-d distribution $\tilde{G}_i$ is a feasible strategy of sender i in the original game. For convenience, for each sender i , we let $\{\tilde{G}^k\}_{k \in \mathbb{Z}^+}$ denote the sequence of discrete equilibrium strategies, and let $\tilde{G}$ denote the limit. We divide this proof into two steps.

First, we prove that $\mathbb{E}_{v \sim \tilde{G}_i}[\max\{v - \theta, 0\}] = c$ for any $\theta \in [-c, 1 - c]$. Let's consider any index $\theta \in (-c, 1 - c)$ since the cases of index $-c$ and index $1 - c$ are trivial. For any $\hat{\theta} \in \text{supp}(\tilde{G}_\theta)$, we assume that

$$\int_{v=0}^1 |c - \max\{v - \theta, 0\}| \tilde{g}(v, \hat{\theta}) dv = d > 0.$$

By the weak convergence of sequence $\{\tilde{G}^k(\hat{v}, \hat{\theta})\}_{k \in \mathbb{Z}^+}$ to distribution $\tilde{G}$, there exists $K_1 > 0$ such that when $k > K_1$, we have

$$\left| \int_{v=0}^1 |c - \max\{v - \theta, 0\}| \tilde{g}(v, \hat{\theta}) dv - \int_{v=0}^1 |c - \max\{v - \theta, 0\}| \tilde{g}^k(v, \hat{\theta}) dv \right| < \frac{d}{2}.$$

This implies that for any $k > K_1$, we have

$$\int_{v=0}^1 |c - \max\{v - \theta, 0\}| \tilde{g}^k(v, \hat{\theta}) dv \in \left(\frac{d}{2}, \frac{3d}{2} \right),$$

¹¹ Take the one-dimensional case as an example, a sequence of functions $\{G^m\}_{m \in \mathbb{Z}^+}$ is tight, if and only if $\forall \epsilon > 0$ there exists an interval $[a, b]$ such that for each $m \in \mathbb{Z}^+$ we have $G^m(b) - G^m(a) > 1 - \epsilon$.

¹² Take the one-dimensional case as an example, sequence $\{G^m\}_{m \in \mathbb{Z}^+}$ weakly converges to CDF G if and only if $\lim_{m \rightarrow \infty} G^m(x) = G(x)$ for each point x at which G is continuous.

which forms a contradiction with the fact $\int_{v=0}^1 |c - \max\{v - \theta, 0\}| \tilde{g}^k(v, \hat{\theta}) dv = 0$, and makes this assumption invalid.

Second, we prove that $\int_{\theta=-c}^{1-c} d\tilde{G}(v, \theta) = f(v)$ for any $v \in [0, 1]$. Let's consider any value $v \in (0, 1]$, since the case of value 0 is trivial. For any $\hat{v} = \alpha \cdot 2^{-\beta}$ for some $\beta \in \mathbb{Z}^+$ and some $\alpha \in [2^\beta]$, then for any $k > \beta$, there exists $t \in \{0\} \cup [2^k]$ such that $\hat{v} = \frac{t}{2^k}$ and $G^k(\frac{t}{2^k}, 1 - c) - G^k(\frac{t-1}{2^k}, 1 - c) = F(\frac{t}{2^k}) - F(\frac{t-1}{2^k})$. Dividing both sides of the equation by 2^{-k} , we get

$$\frac{G^k(\frac{t}{2^k}, 1 - c) - G^k(\frac{t-1}{2^k}, 1 - c)}{2^{-k}} = \frac{F(\frac{t}{2^k}) - F(\frac{t-1}{2^k})}{2^{-k}}.$$

When $k \rightarrow \infty$, we know that $\int_{\theta=-c}^{1-c} dG(\hat{v}, \theta) = f(\hat{v})$. For any $\hat{v} \neq \alpha \cdot 2^{-\beta}$ for any $\beta \in \mathbb{Z}^+$ and any $\alpha \in [2^\beta]$, then for any $k \in \mathbb{Z}^+$ there exist $t \in [2^k]$ such that $\hat{v} \in (\frac{t-1}{2^k}, \frac{t}{2^k})$ and $G^k(\frac{t}{2^k}, 1 - c) - G^k(\hat{v}, 1 - c) = F(\frac{t}{2^k}) - F(\frac{t-1}{2^k})$. Dividing both sides of the equation by 2^{-k} ,

$$\frac{G^k(\frac{t}{2^k}, 1 - c) - G^k(\hat{v}, 1 - c)}{2^{-k}} = \frac{F(\frac{t}{2^k}) - F(\frac{t-1}{2^k})}{2^{-k}}.$$

When $k \rightarrow \infty$, we know that $\int_{-c}^{1-c} dG(\hat{v}, \theta) = f(\hat{v})$.

Combining these two steps, we have proved that each sender i 's strategy in the limit, $\tilde{G}_i$, is indeed a feasible strategy for the original game. $\square$

Step 3: Proving the limit strategy profile $(\tilde{G}_1, \dots, \tilde{G}_N)$ is an equilibrium in the original game. To this end, we need to show that each sender's utility converges in the subsequence along with the convergence of discrete equilibrium strategies. An obstacle to this is that the senders' utilities may be discontinuous in their strategies if there are ties in $(\tilde{G}_1, \dots, \tilde{G}_N)$ with strictly positive probability. To rule out this case, we first prove that, in the subsequence of approximation games, the probability of ties in any neighborhood diminishes. We only need to show this for amortized values that actually affect the senders' utilities (above the smallest winning amortized value defined below).

Definition 7 (Smallest Winning Amortized Value). For any profile of amortized value distributions $(K_1, \dots, K_N)$, the smallest winning amortized value is $\underline{\kappa} \triangleq \max_{i \in [N]} \inf \text{supp}(K_i)$.

Each sender has zero utility for realizing an amortized value below the smallest winning amortized value $\underline{\kappa}$. If sender i 's strategy has a mass in K_i below $\underline{\kappa}$, it does not cause discontinuity in anyone's utility. Therefore we need only to focus on the part of the distribution at or above $\underline{\kappa}$. For the limit strategy profile $(\tilde{G}_1, \dots, \tilde{G}_N)$, denote each sender i 's distribution of amortized value as $\tilde{K}_i$, and the smallest winning amortized value as $\underline{\kappa}$.

Lemma 21. There exist no amortized value $\hat{\kappa} \in [\underline{\kappa}, 1 - c]$ and two distinct senders $i, j \in [N]$ such that $\tilde{K}_i, \tilde{K}_j$ both assign positive probabilities at $\hat{\kappa}$.

Proof of Lemma 21. By the assumption of the prior, there exists no tie at index $1 - c$ in any feasible strategy profile. We assume that there exists an index $\hat{\theta} \in [\underline{\theta}, 1 - c)$ such that sender 1 and 2's strategies, $\tilde{G}_1, \tilde{G}_2$, simultaneously assign a positive mass at index $\hat{\theta}$. Let $\phi_1(\hat{\theta}) - \phi_1(\hat{\theta}^-) = p_1 > 0$ and $\phi_2(\hat{\theta}) - \phi_2(\hat{\theta}^-) = p_2 > 0$. We want to prove that discrete strategies, $\tilde{G}_1^m$ and $\tilde{G}_2^m$, also assign a big enough probability at index $\hat{\theta}$ for sufficiently large m , which violates the equilibrium conditions. By Lemma 20, strategy $\tilde{G}_1$ must be a multi-value row at row $\hat{\theta}$, that is $\inf \text{supp}(\tilde{G}_{1, \cdot | \hat{\theta}}) < \sup \text{supp}(\tilde{G}_{1, \cdot | \hat{\theta}})$. In the same manner, we have $\inf \text{supp}(\tilde{G}_{2, \cdot | \hat{\theta}}) < \sup \text{supp}(\tilde{G}_{2, \cdot | \hat{\theta}})$. First, based on the definition of u , for each sender $i \in [N]$, we define the discrete version of the compressed interim utility as

$$\phi_i^m(x) = \prod_{j \neq i} \left(\tilde{G}_j^m(1, x) + \tilde{G}_j^m(x, 1 - c) - \tilde{G}_j^m(x, x) \right), \quad \forall x \in [-c, 1 - c].$$

Based on whether the index $\hat{\theta}$ belongs to the discrete support of index Θ^m for some $m \in \mathbb{Z}^+$, we divide this part of proof into two cases.

Case 1: There exist $\beta \in \mathbb{Z}^+$ and $\alpha \in [2^\beta]$ such that $\hat{\theta} = \alpha \cdot 2^{-\beta} - c$. We take sender 1 as an example, and the proof of sender 2 follows the same idea. There exists $m_1 > 0$ such that when $m > m_1$, $(\hat{\theta} - 8 \cdot 2^{-m}, \hat{\theta} + 8 \cdot 2^{-m}) \subset (\underline{\theta}, 1 - c)$ and spreading the row $\hat{\theta}$ of the discrete equilibrium strategy to any row $\theta' \in (\hat{\theta}, \hat{\theta} + 8 \cdot 2^{-m})$ and any row $\theta'' \in (\hat{\theta} - 8 \cdot 2^{-m}, \hat{\theta})$ is possible, and will not cause any value of row θ exceed the line $v = \theta$. We define $\epsilon_1 = \frac{p_1}{100}$, then there exists $m_2 > 0$ such that when $m > m_2$, it holds that $|\phi_1^m(\hat{\theta}^+) - \phi_1^m(\hat{\theta}^-) - p_1| < \epsilon_1$. There exists $m_3 > 0$ such that when $m > m_3$, it holds that $\hat{\theta} - 7 \cdot 2^{-m} > \underline{\theta}$ and $\phi_1^m(\hat{\theta}^-) - \phi_1^m(\hat{\theta} - 7 \cdot 2^{-m}) < \epsilon_2$ where $\epsilon_2 = \frac{1}{100}p_1$. There exists $m_4 > 0$ such that when $m > m_4$, it holds that $\hat{\theta} + 2^{-m} < 1 - c$ and $\phi_1^m(\hat{\theta} + 2^{-m}) - \phi_1^m(\hat{\theta}^+) < \epsilon_3$ where $\epsilon_3 = \frac{1}{100}p_1$. Based on these inequalities, we achieve that $3\epsilon_1 + \epsilon_2 < 3p_1$, which implies that

$$\frac{7}{8}(p_1 - \epsilon_1 + \epsilon_2 + \epsilon_3) > \epsilon_2 + \frac{p_1 - \epsilon_1}{2}, \quad \frac{7}{8}(p_1 + \epsilon_1 + \epsilon_2 + \epsilon_3) > \epsilon_2 + \frac{p_1 + \epsilon_1}{2}.$$

This implies that, for any $m > \max\{m_1, m_2, m_3, m_4\}$, spreading the row $\hat{\theta}$ of discrete equilibrium strategy G_1^m to row $\hat{\theta} - 7 \cdot 2^{-m}$ and row $\hat{\theta} + 2^{-m}$ can bring a strictly positive utility increase, which violates the equilibrium conditions.

Case 2: $\hat{\theta} \neq \alpha \cdot 2^{-\beta} - c$ for any $\beta \in \mathbb{Z}^+$ and any $\alpha \in [2^\beta]$. We take sender 1 as an example, and the proof of sender 2 follows the same idea. For any $m \in \mathbb{Z}^+$, there exists $k \in \{0\} \cup [2^m]$ such that $\hat{\theta} \in (\frac{k}{2^m}, \frac{k+1}{2^m})$. There exists $m_1 > 0$ such that when $m > m_1$, it holds that $(k \cdot 2^{-m} - 8 \cdot 2^{-m}, (k+1) \cdot 2^{-m} + 8 \cdot 2^{-m}) \subset (\underline{\theta}, 1 - c)$, and spreading the row $k \cdot 2^{-m}$ of the discrete equilibrium strategy to any row $\theta' \in ((k+1) \cdot 2^{-m}, (k+1) \cdot 2^{-m} + 8 \cdot 2^{-m})$ and any row $\theta'' \in (k \cdot 2^{-m} - 8 \cdot 2^{-m}, k \cdot 2^{-m})$ is possible, and will not cause any value of row $k \cdot 2^{-m}$ exceed the line $v = \theta$. There exists $m_2 > 0$ such that when $m > m_2$, it holds that $|\phi_1^m(((k+1) \cdot 2^{-m})^+) - \phi_1^m((k \cdot 2^{-m})^-) - p_1| < \epsilon_1$ where $\epsilon_1 = \frac{p_1}{100}$. There exists $m_3 > 0$ such that when $m > m_3$, it holds that $k \cdot 2^{-m} - 5 \cdot 2^{-m} > \underline{\theta}$ and $\phi_1^m((k \cdot 2^{-m})^-) - \phi_1^m(k \cdot 2^{-m} - 5 \cdot 2^{-m}) < \epsilon_2$ where $\epsilon_2 = \frac{p_1}{100}$. There exists $m_4 > 0$ such that when $m > m_4$, it holds that $(k+1) \cdot 2^{-m} + 2^{-m} < 1 - c$ and $\phi_1^m((k+1) \cdot 2^{-m} + 2^{-m}) - \phi_1^m(((k+1) \cdot 2^{-m})^+) < \epsilon_3$ where $\epsilon_3 = \frac{p_1}{100}$. Based on these inequalities, we have $\epsilon_1 + 3\epsilon_2 < p_1$, which implies that

$$\frac{5}{8}(p_1 - \epsilon_1 + \epsilon_2 + \epsilon_3) > \epsilon_2 + \frac{p_1 - \epsilon_1}{2}, \quad \frac{5}{8}(p_1 + \epsilon_1 + \epsilon_2 + \epsilon_3) > \epsilon_2 + \frac{p_1 + \epsilon_1}{2}.$$

This implies that, for any $m > \max\{m_1, m_2, m_3, m_4\}$, spreading the row $k \cdot 2^{-m}$ of discrete equilibrium strategy G_1^m to row $k \cdot 2^{-m} - 5 \cdot 2^{-m}$ and row $(k+1) \cdot 2^{-m} + 2^{-m}$ can bring a strictly positive utility increase, which violates the equilibrium conditions.

In summary, given the assumption, the converging discrete strategy profile cannot form an equilibrium, which make this assumption invalid. Thus, we have proved that there is no tie in the limit strategy profile $(\tilde{G}_1, \dots, \tilde{G}_N)$ at and above $\underline{\theta}$. $\square$

With the obstacle of discontinuity cleared, we obtain the convergence of utility functions along with the convergence of discrete equilibrium strategies. Here we comes the final step of the proof.

Lemma 22. *The limit profile $(\tilde{G}_1, \dots, \tilde{G}_N)$ is indeed an equilibrium in the original game.*

Proof of Lemma 22. Consider each sender $i \in [N]$ and any feasible strategy $G_i \in \mathcal{G}_i(F_i, c)$, we want to show that strategy $\tilde{G}_i$ achieves a higher expected utility than strategy G_i , given others strategies $\tilde{G}_{-i}$. We divide this part of the proof into two cases.

Case 1: There exist ties over $[\underline{\theta}, 1 - c]$ in strategy profile $(G_i, \tilde{G}_{-i})$. We assume that there exists a tie at index $\hat{\theta}$ in strategy profile $(G_i, \tilde{G}_{-i})$. By the assumption of the prior, strategy G_i

must be a multi-value row at row $\hat{\theta}$, that is $\inf \text{supp}(G_{i, \cdot | \hat{\theta}}) < \sup \text{supp}(G_{i, \cdot | \hat{\theta}})$, which implies that strategy G_i can spread the probability of index $\hat{\theta}$ to index $\hat{\theta} + \epsilon$ and index $\hat{\theta} - \epsilon$ for sufficiently small $\epsilon > 0$. By the fact that $\phi_i(\hat{\theta}^-) < \phi_i(\hat{\theta}) < \phi_i(\hat{\theta}^+)$, we know spreading the probability of strategy G_i at index $\hat{\theta}$ to index $\hat{\theta} + \epsilon$ and index $\hat{\theta} - \epsilon$ achieves a utility increase for sufficiently small $\epsilon > 0$. Therefore, we know that, given others' strategies, the best response strategy of sender i will not form a tie with the strategies of others within the interval $[\underline{\theta}, 1 - c]$. Thus, in Case 2, we only need to prove that for those strategies that do not form a tie with the others' strategies within $[\underline{\theta}, 1 - c]$, strategy $\tilde{G}_i$ will achieve a higher expected utility. This implies that strategy $\tilde{G}_i$ is the best response to others' strategies.

Case 2: There is no tie over $[\underline{\theta}, 1 - c]$ in strategy profile $(G_i, \tilde{G}_{-i})$. Let $\tilde{u}_i$ denote sender i 's interim utility function given others' strategies $\tilde{G}_{-i}$. Let $\tilde{u}_i^m$ denote sender i 's interim utility function given others' strategies $\tilde{G}_{-i}^m$. By Lemma 21, there is no tie over $[\underline{\theta}, 1 - c]$ in the limit strategy profile $(\tilde{G}_1, \dots, \tilde{G}_N)$. By the definition of weak convergence, we have that

$$\lim_{m \rightarrow \infty} \int_{[0,1] \times [-c, 1-c]} \tilde{u}_i^m(v, \theta) \tilde{g}_i^m(v, \theta) \, dv \, d\theta = \int_{[0,1] \times [-c, 1-c]} \tilde{u}_i(v, \theta) \tilde{g}_i(v, \theta) \, dv \, d\theta. \quad (23)$$

For each sender $i \in [N]$ and any feasible strategy $G_i \in \mathcal{G}_i(F_i, c)$ that do not form a tie with others' strategies within $[\underline{\theta}, 1 - c]$, we construct a 2-dimensional CDFs sequence $\{G_i^m\}_{m \in \mathbb{Z}^+}$ where each $G_i^m \in \mathcal{S}_i^m$, and the sequence weakly converges to strategy G_i . Specifically, for $\forall m \in \mathbb{Z}^+$, we define G_i^m as below.

$$G_i^m(v, \theta) = \begin{cases} G_i(v, \theta), & \text{if } (v, \theta) \in V^m \times \Theta^m, \\ G_i(\max\{t : t \in V^m, t \leq v\}, \max\{t : t \in \Theta^m, t \leq \theta\}), & \text{otherwise.} \end{cases}$$

First, we show that the sequence $\{G_i^m\}_{m \in \mathbb{Z}^+}$ weakly converges to strategy G_i . We consider any pair of $(v, \theta) \in [0, 1] \times [-c, 1 - c]$. If there exist $\alpha_1 \in \mathbb{Z}^+$ and $\beta_1 \in [2^{-\alpha_1}]$ such that $v = \beta_1 \cdot 2^{-\alpha_1}$ and $\alpha_2 \in \mathbb{Z}^+$ and $\beta_2 \in [2^{-\alpha_2}]$ such that $\theta = \beta_2 \cdot 2^{-\alpha_2}$, then when $m > \max\{\alpha_1, \alpha_2\}$, we have $G_i^m(v, \theta) = G_i(v, \theta)$ which implies that sequence $\{G_i^m(v, \theta)\}_{m \in \mathbb{Z}^+}$ converges to $G_i(v, \theta)$. If $v \neq \beta_1 \cdot 2^{-\alpha_1}$ for any $\alpha_1 \in \mathbb{Z}^+$, $\beta_1 \in [2^{-\alpha_1}]$ or $\theta \neq \beta_2 \cdot 2^{-\alpha_2}$ for any $\alpha_2 \in \mathbb{Z}^+$, $\beta_2 \in [2^{-\alpha_2}]$, then for any $m \in \mathbb{Z}^+$, we have $G_i^m(v, \theta) = G_i(\max\{t : t \in V^m, t \leq v\}, \max\{t : t \in \Theta^m, t \leq \theta\})$. When m goes to infinity, sequence $\{\max\{t : t \in V^m, t \leq v\}\}_{m \in \mathbb{Z}^+}$ converges to v , and sequence $\{\max\{t : t \in \Theta^m, t \leq \theta\}\}_{m \in \mathbb{Z}^+}$ converges to θ . So we have sequence $\{G_i^m(v, \theta)\}_{m \in \mathbb{Z}^+}$ also converges to $G_i(v, \theta)$. By the fact that sequence $\{G_i^m\}_{m \in \mathbb{Z}^+}$ weakly converges to strategy G_i , using a similar method of Lemma 20, we can also prove that there exists $m_1 > 0$ such that when $m < m_1$, $G_i^m \in \mathcal{S}_i^m$.

By Equation (23), we have

$$\begin{aligned} & \int_{[0,1] \times [-c, 1-c]} \tilde{u}_i(v, \theta) g_i(v, \theta) \, dv \, d\theta - \int_{[0,1] \times [-c, 1-c]} \tilde{u}_i(v, \theta) \tilde{g}_i(v, \theta) \, dv \, d\theta \\ &= \lim_{m \rightarrow \infty} \left(\int_{[0,1] \times [-c, 1-c]} \tilde{u}_i(v, \theta) g_i(v, \theta) \, dv \, d\theta - \int_{[0,1] \times [-c, 1-c]} \tilde{u}_i^m(v, \theta) g_i^m(v, \theta) \, dv \, d\theta \right) \\ &+ \lim_{m \rightarrow \infty} \left(\int_{[0,1] \times [-c, 1-c]} \tilde{u}_i^m(v, \theta) g_i^m(v, \theta) \, dv \, d\theta - \int_{[0,1] \times [-c, 1-c]} \tilde{u}_i^m(v, \theta) \tilde{g}_i^m(v, \theta) \, dv \, d\theta \right) \\ &= \lim_{m \rightarrow \infty} \left(\int_{[0,1] \times [-c, 1-c]} G_i^m(v, \theta) \, d\tilde{u}_i^m(v, \theta) - \int_{[0,1] \times [-c, 1-c]} G_i(v, \theta) \, d\tilde{u}_i(v, \theta) \right) \\ &+ \lim_{m \rightarrow \infty} \left(\int_{[0,1] \times [-c, 1-c]} \tilde{u}_i^m(v, \theta) g_i^m(v, \theta) \, dv \, d\theta - \int_{[0,1] \times [-c, 1-c]} \tilde{u}_i^m(v, \theta) \tilde{g}_i^m(v, \theta) \, dv \, d\theta \right). \end{aligned}$$

Combining the facts that sequence $\{(\tilde{G}_1^m, \dots, \tilde{G}_N^m)\}_{m \in \mathbb{Z}^+}$ weakly converges to $(\tilde{G}_1, \dots, \tilde{G}_N)$, sequence $\{G_i^m\}_{m \in \mathbb{Z}^+}$ weakly converges to G_i , and there is no tie over $[\theta, 1 - c]$ in strategy profile $(G_i, \tilde{G}_{-i})$, we have that

$$\lim_{m \rightarrow \infty} \left(\int_{[0,1] \times [-c, 1-c]} G_i(v, \theta) du_i(v, \theta) - \int_{[0,1] \times [-c, 1-c]} G_i^m(v, \theta) du_i^m(v, \theta) \right) = 0. \quad (24)$$

In addition to the fact $(\tilde{G}_1^m, \dots, \tilde{G}_N^m)$ is an equilibrium in the m -th discrete approximation game, we have

$$\lim_{m \rightarrow \infty} \left(\int_{[0,1] \times [-c, 1-c]} u_i^m(v, \theta) g_i^m(v, \theta) dv d\theta - \int_{[0,1] \times [-c, 1-c]} u_i^m(v, \theta) \tilde{g}_i^m(v, \theta) dv d\theta \right) \leq 0. \quad (25)$$

Combining Inequalities (24) and (25), we have

$$\int_{[0,1] \times [-c, 1-c]} u_i(v, \theta) g_i(v, \theta) dv d\theta - \int_{[0,1] \times [-c, 1-c]} u_i(v, \theta) \tilde{g}_i(v, \theta) dv d\theta \leq 0,$$

which shows that for each sender i , strategy $\tilde{G}_i$ is a best response to others' strategies $\tilde{G}_{-i}$, and $(\tilde{G}_1, \dots, \tilde{G}_N)$ is indeed an equilibrium in our game. $\square$

Putting all the pieces together, finally we can prove Theorem 5.

Proof of Theorem 5. Theorem 5 holds directly by combining Lemmas 18 to 22. $\square$

E Symmetric Equilibrium under Concave Priors

Here we show that the characterization of symmetric equilibrium of concave priors is much more involved than that of convex priors. We consider a two-sender game where $\text{supp}(F) = [c, 1]$, and the prior F is concave over its support. Let $S = \int_0^1 F(x) dx$, and we know $S \geq \frac{1}{2}$ since the concavity of the prior.

Different values of the cost c lead to distinct symmetric equilibrium structures, making it challenging to unify all cases under a single amortized value distribution or a generalized 2-D distribution construction. While a complete equilibrium characterization remains an open problem, we present partial results below, specifically detailing the equilibrium structure for cost regimes within certain intervals.

When the cost c is sufficiently large, we make the following construction of a simplest kind of symmetric equilibrium.

Theorem 6. *When the cost $c \geq \frac{2}{3}(1 - S)$, there exists a symmetric equilibrium with the amortized value distribution $K(\cdot)$ defined as below:*

$$K(\theta) = \begin{cases} F(\theta + c) & \text{if } \theta \in [-c, 0], \\ \min \left\{ \frac{1}{\bar{\theta}} \theta, 1 \right\} & \text{if } \theta \in (0, 1 - c], \end{cases}$$

where $\bar{\theta}$ is the unique solution to the equation $\int_0^{1-c} \min \left\{ \frac{1}{\bar{\theta}} \theta, 1 \right\} d\theta = \int_0^{1-c} F(\theta + c) d\theta$.

Proof of Theorem 6. This theorem follows a similar proof of Theorem 2. The construction method can be seen in Figure 7. $\square$

This type of symmetric equilibrium may fail to exist when $c < \frac{2}{3}(1 - S)$, since in that case it holds $\bar{\theta} > c$, our construction of 2-D distribution may break down halfway. Although this amortized value distribution forms a candidate for equilibrium, it cannot be guaranteed that there exists a 2-D distribution that induces this amortized value distribution.

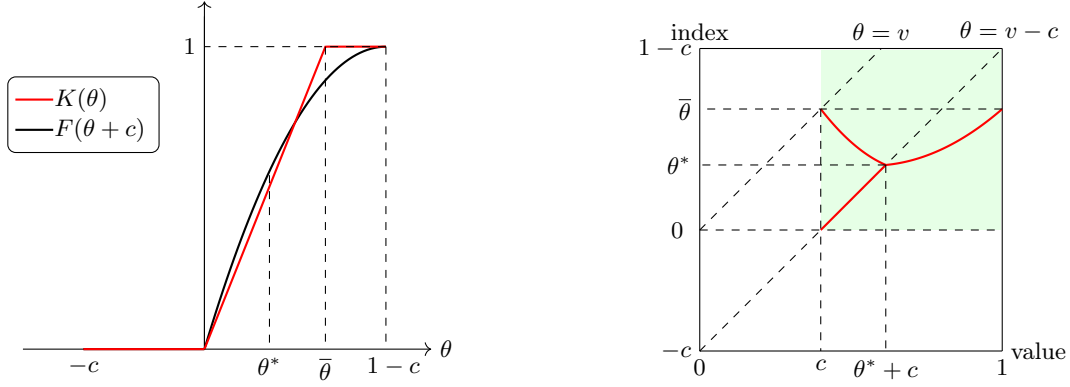


Fig. 7. An example of Theorem 6. In the left panel, the red curve represents the amortized value distribution $K(\cdot)$, and the black curve represents the shifted prior. Function $K(\cdot)$ forms an MPC of the shifted prior over $[-c, 1-c]$. In the right panel, the red curves represent the support set of the corresponding 2-D distribution that forms an equilibrium.

F Setting Comparison of the Competitive Information Design Game

As mentioned in Related Works, Ding et al. (2023); Hwang and Hwang (2025) study a closely related model, where the sender’s signal is seen by the buyer *after* inspection, and the buyer sees only the signal instead of her value. We call this setting one of *information obfuscation*; in comparison, our setting is one of *information revelation*. In this section, we first compare the informational properties of these two settings and then compare the senders’ and the receiver’s utility.

F.1 Setting Comparison

Extreme cases under both settings. Extreme cases highlight the differences between the two models. In our model, all senders adopting the *full revelation* strategy let the buyer know all the values before searching; her utility would be the first best $\mathbb{E}[\max_i (v_i - c_i)_+]$. If no sender sends any signal, the buyer gets no additional information. This reduces the case to the original Pandora Box problem where the buyer just searches with the original priors $F_1, \dots, F_N$. In contrast, this original problem corresponds to the case of full revelation in the setting of information obfuscation. In this setting, if the senders send no signals, the distribution of each v_i degenerates to a point mass on $\mathbb{E}[v_i]$, which reduces the buyer’s utility to $\max_i (\mathbb{E}[v_i] - c_i)_+$.

Buyer’s search behavior under both settings. The buyer’s search behavior also differs in the two models. In the information revelation setting, the buyer’s search depends on the *signals* she receives, updating her posteriors based on these signals. In the obfuscation setting, the search depends only on the *signaling schemes* — once the senders commit to the schemes, the value distribution of each seller becomes an MPC of the original prior, and the receiver searches using these MPCs, observing a signal only after inspecting a seller.

F.2 Utility Comparison

Receiver Utility Comparison. We treat the original Pandora Box problem as the benchmark. Fixing all boxes priors and search costs, let U be the buyer’s utility when she implements the Index Algorithm in the original Pandora Box problem.

Proposition 2 (Receiver Utility Comparison). *U is weakly lower than the buyer’s utility in the information revelation setting, and weakly higher than the buyer’s utility in the information obfuscation setting, regardless of the senders’ signaling schemes in both settings.*

Proof of Proposition 2. Under the information revelation setting, one search strategy available to the receiver is to ignore the signals and to implement the Index Algorithm based on the original priors. This searching strategy yields the same overall utility as in the no information setting. But this utility is no more than that of the Index Algorithm performed on the posteriors, since the posteriors are more informative than the priors and the agent can better estimate the value based on the posteriors than based on the priors. This proves the first claim.

For the second claim, observe that in the no information setting, the receiver can simulate the obfuscation setting by intentionally ignoring the true value v_i of any inspected box i and instead behaving as if she had observed its posterior mean value. The utility of this simulation is precisely that in the obfuscation setting, which is no more than the utility of the Index Algorithm on the original priors, due to the optimality of Index Algorithm. This proves the second claim. $\square$

The second statement was first shown by [Ding et al. \(2023\)](#) (Theorem 3.1); we consider our proof here considerably simpler. Intuitively, the more information accessible to the receiver, the higher her utility should be.

Sender Utility Comparison. The next proposition considers a sender choosing between revealing or obfuscating information.

Proposition 3 (Sender Utility Comparison). *Fix all priors and costs, and consider a sender i who unilaterally considers deploying a signaling scheme, we have:*

- (i) *when $c_i = 0$, any strategy of obfuscation gives him weakly higher utility than revelation;*
- (ii) *when $c_i > \mathbb{E}_{v_i \sim F_i}[v_i]$, any strategy of revelation gives him weakly higher utility than obfuscation.*

Proof of Proposition 3. When the inspection cost is zero for sender i , the receiver does any inspection for free. Thus, it is optimal for her to inspect all boxes and chooses the most favorable one. An information revealing sender can do nothing, since any information revelation only serves to influence the searching order of the agent. While for an obfuscating sender, he can influence the final choice of the agent through certain obfuscation strategy. Therefore, an obfuscating sender has weakly higher utility.

When the inspection cost is higher than the expectation of the prior distribution, an obfuscating sender always has a negative index regardless of any obfuscation strategy. Thus, he can do nothing in this case, since any box with a negative index will never be inspected under the Index Algorithm. On the other hand, an information-revealing sender can sacrifice some low values and bundle higher values together to form strictly positive indices, thereby ensuring an expected payoff greater than zero. Therefore the latter has a weakly higher utility. $\square$

The cost affects the receiver's willingness to search and her optimal search strategy. For a seller i with search cost $c_i = 0$, the receiver will definitely inspect the seller for free. Thus, an information-revealing sender can do nothing, since any revelation only serves to influence the searching order of the receiver. In contrast, when his cost is sufficiently high, an obfuscating sender always has a negative index regardless of the strategy, thus will never be inspected under the Index Algorithm.