

Information Design in Optimal Auctions*

Yi-Chun Chen[†]

Xiangqian Yang[‡]

October 15, 2022

Abstract

We study the information design problem in a single-unit auction setting. The information designer controls independent private signals according to which the buyers infer their binary private values. Assuming that the seller adopts the optimal auction due to [Myerson \(1981\)](#) in response, we characterize both the buyer-optimal information structure, which maximizes the buyers' surplus, and the seller-worst information structure, which minimizes the seller's revenue. We translate both information design problems into finite-dimensional, constrained optimization problems in which one can explicitly solve for the optimal information structure. In contrast to the case with one buyer ([Roesler and Szentes, 2017](#)), we show that with two or more buyers, the symmetric buyer-optimal information structure is different from the symmetric seller-worst information structure. The good is always sold under the seller-worst information structure but not under the buyer-optimal information structure. Nevertheless, as the number of buyers goes to infinity, both symmetric information structures converge to no disclosure. We also show that in our ex ante symmetric setting, an asymmetric information structure is never seller-worst but can generate a strictly higher surplus for the buyers than the symmetric buyer-optimal information structure.

Keywords— information design; optimal auction; virtual value distribution; buyer-optimal information; seller-worst information.

*We are very grateful to the editor and two anonymous referees for their valuable comments. We also thank Soumen Banerjee, Ben Brooks, Songzi Du, Simone Galperti, Gaoji Hu, Jingfeng Lu, Anne-Katrin Roesler, Satoru Takahashi, Kai Hao Yang, Wanchang Zhang, Xinhan Zhang, Weijie Zhong, and Junjie Zhou for their valuable comments. Financial support from “the Singapore Ministry of Education Academic Research Fund Tier 1” is gratefully acknowledged. All errors are our own.

[†]Department of Economics and Risk Management Institute, National University of Singapore. Email: ecsycc@nus.edu.sg

[‡]School of Economics and Trade, Hunan University, China. Email: yangxiangqian@hnu.edu.cn

1 Introduction

Consider a seller who would like to sell one object to a group of buyers. The classical optimal auction due to [Myerson \(1981\)](#) assumes that each buyer privately knows his own valuation; moreover, each valuation follows a distribution which is common knowledge. In this paper, we study an information design problem in which each buyer learns his private valuation independently via a signal according to which the seller runs the Myersonian optimal auction. The seller earns the expected highest nonnegative virtual value, whereas the buyers earn the expected total surplus minus the seller’s revenue. We derive *the buyer-optimal information structure* which maximizes the buyers’ total surplus, as well as *the seller-worst information structure* which minimizes the seller’s optimal revenue.

In reality the buyers may not know their own valuations for the good and have to assess how well the product suits their need via information sources such as advertisements, recommendations from some platform, or product descriptions. For instance, personalized advertising communicates privately with the buyers according to their individual characteristics such as gender, age, economic status, and so on. We consider contexts in which these personal characteristics are independently distributed so that the information is purely private,¹ that is, one learns nothing about a buyer’s information or characteristics from the advertisement shown to another buyer. These features motivate our study of information design with independent private signals.

Providing more information to the buyers can lead to a higher surplus but also a higher payment for them in an optimal auction. Hence, the effect of a new information source on the buyers’ welfare is not a priori clear. The buyer-optimal information structure contributes to our understanding of this issue by identifying an information structure which maximizes the buyers’ aggregate surplus. In this regard, our study builds upon the prior work on monopoly pricing by [Roesler and Szentes \(2017\)](#) but expands the scope to an auction setup with multiple buyers. The information designer may be a regulator who aims to promote consumers’ welfare by requiring the seller to disclose certain information about the product, or else by restricting the seller from doing so.² The information designer may also be a data vendor who can sell product-related information for a fee proportional to buyers’ (average/ total) surplus and who therefore looks for an information structure which maximizes the buyers’ surplus.

The buyers’ surplus is the total surplus minus the seller’s revenue. Hence, studying the seller-worst information structure helps us understand the trade-off between minimizing the seller’s revenue and maximizing the total surplus. The seller-worst information design also offers a “minmax” upper bound

¹See, for instance, [Chen, Owen, Pixton, and Simchi-Levi \(2022\)](#) where the consumers’ characteristics are modeled as i.i.d. random vectors.

²[Terstiege and Wasser \(2020\)](#) study a buyer-optimal information design problem with monopoly pricing in which the information designer may be a regulator of the product information. For example, prescription drug ads must list side-effects and contra-indications to protect consumers. They focus on a situation where the buyer cannot commit to ignore any additional information released by the seller, whereas we follow [Roesler and Szentes \(2017\)](#) in setting aside the issue of the seller’s disclosure.

for the *revenue guarantee* of a mechanism regardless of the equilibrium and the information structure. Such a “minmax” upper bound is crucial to establishing the strong duality results in [Bergemann, Brooks, and Morris \(2016\)](#), [Du \(2018\)](#), and [Brooks and Du \(2021\)](#). Specifically, these strong duality results show, in different contexts, that this “minmax” upper bound is equal to the maximal revenue guarantee achieved by a “maxmin” mechanism. In this vein, the seller-worst information provides a first step toward establishing such a strong duality result—or lack thereof, in an independent private-value setting.³

We assume that the seller has zero reservation value for the good and the buyers are ex ante symmetric. In particular, the buyers’ ex post valuation of the good is equal to either 0 or 1 with an identical mean p . We assume that each signal provides an unbiased estimator about the buyer’s valuation. Hence, by [Blackwell \(1953\)](#), an information structure is feasible if and only if it consists of a profile of independent signal distributions, all with mean p .⁴ We begin by solving an optimal symmetric signal distribution in both information design problems.⁵ As is true for deriving symmetric equilibria in symmetric auctions, it is also more tractable to derive a symmetric information structure in our ex ante symmetric information design problems.⁶

We show that as long as there are two or more buyers, the buyer-optimal information structure need not be equal to the seller-worst information structure. This result sharply contrasts with the results of [Roesler and Szentes \(2017\)](#), which show that the two information structures are equivalent when there is only one buyer. More precisely, when there are two or more buyers, we pin down a cutoff p_s which is decreasing with the number of buyers. If p is no more than p_s , then the (symmetric) seller-worst information structure for each buyer remains the same as in the one-buyer case. If p is higher than p_s , then the seller-worst signal distribution remains equal to a truncated Pareto distribution but with virtual value $k_s > 0$ for any signal less than 1, and with virtual value 1 when the signal is equal to 1. Since

³We discuss the tightness of this upper bound in Section 6.3. In particular, [Bachrach, Chen, Talgam-Cohen, Yang, and Zhang \(2022\)](#) recently prove that when there are only two buyers, a second-price auction with a suitably chosen random reserve price guarantees exactly the seller-worst revenue that we identify in this paper over all symmetric independent information structures and undominated equilibria.

⁴Alternatively, our analysis also applies and produces the same result if the information designer is allowed to choose any signal distributions with a given mean p and support $[0, 1]$. In a similar vein, [Carrasco, Luz, Kos, Messner, Monteiro, and Moreira \(2018\)](#) study a revenue-maximizing seller with a single buyer where the seller has only partial information about the buyer’s valuation distribution, e.g., its first and second moments.

⁵Focusing on the seller-optimal information, [Bergemann and Pesendorfer \(2007\)](#) considers a Myersonian optimal auction setting where the seller can decide to whom to sell at what price and the accuracy by which bidders learn their (not necessarily binary) valuation through independent private signals. In our setting, the seller-optimal information is full revelation (i.e., the prior) against which the seller posts the price 1 and extracts full surplus. This is a special case of the full-surplus-extraction result due to [Krähmer \(2020\)](#).

⁶Even with symmetric information structures, we still allow for irregular signal distributions for which the optimal auction need not be a second-price auction with a reserve price. Hence, our information design problem is not equivalent to the corresponding information design problem where the seller is committed to adopting a second-price auction with reserve; see Section 4.4 for more discussions and Appendix A.12 for an illustrative example.

all virtual values are nonnegative at any signal profile, the good is always sold. Indeed, raising the low virtual value from 0 to k_s has two countervailing effects. First, the seller's revenue increases as the low virtual value increases. Second, to satisfy the mean constraint, increasing the low virtual value must be compensated for by decreasing the probability of having the high virtual value 1. Through a cost-benefit analysis, we show that the low virtual value k_s increases when either the prior mean or the number of buyers grows.

The (symmetric) buyer-optimal information structure differs from the (symmetric) seller-worst information structure in several ways. First, we pin down two cutoffs r_b and p_b which are also decreasing in the number of buyers. If p lies between r_b and p_b , then the buyer-optimal information structure remains the same as in the one-buyer case. When p is less than r_b , the buyer-optimal signal distribution puts a positive mass on signal 0 and the remaining mass on a truncated Pareto distribution with virtual values 0 and 1. Since signal 0 induces a negative virtual value, with positive probability the seller withholds the good. The reason for this is that, to maximize the buyers' surplus, the information designer needs to consider not only the seller's revenue but also the total surplus. Moreover, the total surplus is convex in the buyers' signal and hence favors dispersion. When p is above p_b , the buyer-optimal signal distribution is a truncated Pareto distribution. However, also due to the convexity of the total surplus, the distribution induces a low virtual value $k_b < k_s$ for any signal less than 1 and a high virtual value 1 otherwise.

Our results highlight two kinds of allocative inefficiency with multiple buyers. First, in our setting the good will be allocated to a buyer with the highest interim (virtual) value, who need not have the highest ex post value 1. Second, as we argue above, the buyer-optimal signal distribution may put a positive mass on signal 0 against which the seller withholds the good. Both sorts of inefficiency sharply contrast with the optimal information with one buyer under which ex post efficiency is always achieved.

When the number of buyers goes to infinity, the cutoffs p_s , r_b , and p_b all tend to zero; both k_b and k_s monotonically increase to p ; and the corresponding probabilities assigned to virtual value 1 monotonically decrease. As a result, both the buyer-optimal and the seller-worst signal distributions converge to a degenerate distribution which puts all mass on the prior mean p . In particular, learning *no* information is asymptotically buyer-optimal with a large number of buyers. This result offers an extreme form of the message due to [Roesler and Szentes \(2017\)](#) that a buyer does not want to learn his valuation perfectly in a monopoly setting. This result also sharply contrasts with the result of [Yang \(2018\)](#) that when the buyers engage in strategic information acquisition, they will acquire (in a symmetric equilibrium) asymptotically perfect information about their values.

We summarize the results on optimal symmetric information in Table 1.

We also investigate asymmetric information structure in both information design problems. For the seller-worst problem, we show that the optimal symmetric information structure remains the unique optimal solution, even if the information designer can choose different signal distributions for different

Table 1: Main results and comparisons

Number of buyers	Optimal distribution with different prior mean p				Always sell?
$n = 1$	Seller-worst	Virtual values 0 and 1 for any prior mean p , which is consistent with Roesler and Szentes (2017)			Yes
	Buyer-optimal				
$n \geq 2$	Seller-worst	$0 < p \leq p_s$		$p_s < p < 1$	Yes
		Virtual values 0 and 1		Virtual values k_s and 1	
	Buyer-optimal	$0 < p < r_b$	$r_b \leq p \leq p_b$	$p_b < p < 1$	No
		A positive mass on 0 as well as virtual values 0 and 1	Virtual values 0 and 1	Virtual values k_b and 1	
$n \rightarrow \infty$	Seller-worst	Degenerate distribution (revealing nothing) for any mean			Yes
	Buyer-optimal				

buyers. Intuitively, averaging a profile of asymmetric virtual value distributions grants the seller fewer option values in selecting the highest virtual values and thereby earns him less revenue. This means that restricting attention to symmetric signal distributions entails no loss in minimizing the seller's revenue. However, averaging a profile of asymmetric signal distributions may entail loss in the expected total surplus. In particular, we explore one case with two buyers and another case with a large number of buyers, in both of which an asymmetric information structure generates a strictly higher aggregate surplus for the buyers than the optimal symmetric information structure does. Our result shows that asymmetric information structure can emerge endogenously as a choice of a buyer-optimal information designer, even in our ex ante symmetric setting.

Finally, we explain the novelty of our argument. For us to solve our information design problems, it is crucial that we transform the control variables. More precisely, instead of working with signal/interim value distributions, we work with the interim *virtual* value distribution. After the change of variable, the information design problem becomes an isoperimetric problem in optimal control theory. The Euler-Lagrange equation can then be invoked in this problem to argue that the virtual value distribution function is a step function with at most two steps. This effectively reduces the infinite-dimensional information design problem to a tractable finite-dimensional constrained optimization problem. Moreover, with the few control variables, such as the two-step virtual value distribution functions, we are able to understand their trade-off, as we have explained above.

The rest of this paper proceeds as follows. Section 2 describes our model and formulates the information design problem. Section 3 presents our main results. Section 4 demonstrates how we simplify the control variables of the information design problems. Section 5 studies the information design problem with asymmetric signal distributions. Section 6 discusses issues with asymmetric or continuous priors

and studies the tightness of our seller-worst upper bound for some candidate “maxmin” mechanisms. Section 7 concludes. Appendix A contains all proofs which are omitted from the main text.

2 Model

There is a seller who has one object to sell to a finite set $N = \{1, 2, \dots, n\}$ of potential buyers. The seller has no value for the object. Each buyer’s prior valuation, v_i , is identically and independently drawn from a Bernoulli distribution H on $\{0, 1\}$. Let $p = \mathbb{E}[v_i] = \Pr(v_i = 1)$ denote the mean of H . To rule out trivial cases, we assume that $p \in (0, 1)$. Suppose that (i) each buyer can observe an independently and identically distributed signal x_i about v_i from an information designer, and (ii) the joint distribution of v_i and x_i is common knowledge to the seller as well as among the buyers. That is, the information designer commits to a signal structure for each agent, and each agent observes that commitment, not just for his own signal structure but also for the other agents’ signal structures.

2.1 Information structure

Following [Roesler and Szentes \(2017\)](#), we say a signal distribution is *feasible* if each signal of a buyer provides him with an unbiased estimate about his valuation. Then, according to the characterization of [Blackwell \(1953\)](#), the prior valuation distribution H is a mean-preserving spread of any feasible distribution of signals. Since H is a Bernoulli distribution on $\{0, 1\}$, the mean-preserving spread condition can be reduced to a mean constraint. Hence, a feasible symmetric information structure is a signal distribution G with $G \in \mathcal{G}_H$, where

$$\mathcal{G}_H = \left\{ G : [0, 1] \rightarrow [0, 1] \left| \int_0^1 x \, dG(x) = p \text{ and } G \text{ is a CDF} \right. \right\}.$$

2.2 Information design problem

Given a feasible signal distribution G , a revenue-maximizing mechanism is an optimal auction due to [Myerson \(1981\)](#). In an optimal auction, the seller’s revenue is equal to the expected highest, nonnegative, ironed virtual value $\max_i \{\hat{\varphi}(x_i|G), 0\}$. Formally, for any CDF G with $\text{supp}(G) \subset [0, 1]$, let $a = \inf\{x \in [0, 1] | G(x) > 0\}$, and define

$$\Psi(x|G) = \begin{cases} 0, & \text{if } x \in [0, a); \\ a - x(1 - G(x)), & \text{if } x \in [a, 1]. \end{cases}$$

Let $\Phi(x|G)$ be the convexification of Ψ under the G -quantile space.⁷ By definition, for any $x \in [0, 1]$, the (ironed) virtual valuation at x , denoted as $\hat{\varphi}(x|G)$, is an infimum of the G -sub-gradients of $\Phi(x|G)$.

⁷That is, $\Phi(x|G)$ is the largest convex function of $G(x)$ that is everywhere weakly below $\Psi(x|G)$. In [Appendix A.1](#), we also give a formal and detailed instruction about the ironed virtual value from [Monteiro and Svaiter \(2010\)](#) and also [Yang \(2018\)](#).

If $\Phi(x|G) = \Psi(x|G)$ for any x , then we say that G is a *regular distribution*. We denote by $\hat{\varphi}$ an ironed virtual value and use φ to denote a virtual value induced from a regular distribution. If G is regular, then the virtual value has the well-known expression equal to

$$\varphi(x|G) = x - \frac{1 - G(x)}{G'(x)}.$$

We allow the information designer to choose any feasible distribution function G , whether it is regular or irregular and whether or not it admits a density function.

Let $M(x|G) = \{i \in N | \hat{\varphi}(x_i|G) \geq \max_j \{\hat{\varphi}(x_j|G), 0\}\}$ be the set of buyers who have the largest nonnegative virtual value for a given signal realization x ; and let $M'(x|G) = \{i \in N | x_i \geq x_j, \forall j \in M(x|G)\}$ be the set of buyers who have not only the highest nonnegative virtual value but also the largest signal among those with the highest virtual value for a given signal realization x . Define an allocation rule as follows:

$$q_i(x_i, x_{-i}|G) = \begin{cases} \frac{1}{|M'(x|G)|}, & \text{if } i \in M'(x|G); \\ 0, & \text{if } i \notin M'(x|G). \end{cases}$$

That is, $q_i(x_i, x_{-i}|G)$ is an optimal auction allocation rule which breaks a tie in favor of surplus maximization.

We study the following information design problem parameterized by $\alpha = 0$ or 1 :

$$\max_G \int_{[0,1]^n} \sum_{i=1}^n (\alpha x_i - \hat{\varphi}(x_i|G)) q_i(x_i, x_{-i}|G) \prod_{i=1}^n (dG(x_i)) \quad (1)$$

$$\text{s.t. } \int_0^1 (1 - G(x)) dx = p. \quad (2)$$

The term $\sum_{i=1}^n x_i q_i(x_i, x_{-i}|G)$ is the total surplus generated under the optimal auction allocation rule q_i . Moreover, the term $\sum_{i=1}^n \hat{\varphi}(x_i|G) q_i(x_i, x_{-i}|G)$ is the seller's revenue under the allocation rule q_i , namely, the expected highest nonnegative virtual value. Hence, if $\alpha = 0$, the information designer aims to minimize the seller's revenue, and this corresponds to the seller-worst information design problem. If $\alpha = 1$, the information designer aims to maximize the buyers' surplus, and this corresponds to the buyer-optimal information design problem. Hereafter, we call (2) the *mean constraint*.

Endow the space of Borel probability measures on $[0, 1]$ with the weak* topology. We say a signal distribution G induces *nonnegative virtual values except at 0* if the virtual values induced by G are nonnegative almost everywhere on $(0, 1]$. We denote by $\mathcal{G}_H^+ \subset \mathcal{G}_H$ the feasible signal distribution with nonnegative virtual values except at 0. Also we say a signal distribution G is *regular except at 0* if G is regular almost everywhere on $(0, 1]$. We will argue later in Lemma 4 and in Lemma 5, the optimal signal distribution must induce nonnegative virtual values except at 0 and be regular except at 0. We now present two lemmas. The first lemma documents the existence of the solution to the problem in (1). The second lemma highlights the additional trade-off which a buyer-optimal information designer is facing, on top of minimizing the seller's revenue.

Lemma 1. *For the problem in (1), an optimal solution exists.*

Proof. We first establish the seller-worst case. By Theorem 2 of Monteiro (2015), the expected revenue is a lower semicontinuous function in G . Hence, the objective function of the problem in (1) is an upper semicontinuous function in G . Since $\mathcal{G}_H$ is a closed subset of the set of Borel probability measures on $[0, 1]$, $\mathcal{G}_H$ is compact. Thus, by the extreme value theorem, an optimal solution exists.

For the buyer-optimal problem, proving the existence of an optimal solution is more involved; we provide a formal proof in Appendix A.2. $\square$

Lemma 2. *For the problem in (1) and signal distribution $\hat{G} \in \mathcal{G}_H^+$, if $\hat{G}$ is a mean-preserving spread of signal distribution G , then $\hat{G}$ will generate more total surplus than G ; and if $\hat{G}$ is a strict mean-preserving spread of signal distribution G , then $\hat{G}$ will strictly generate more total surplus than G .⁸*

Proof. Observe that $\sum_{i=1}^n x_i q_i(x_i, x_{-i}|G) = \max\{x_1, \dots, x_n\}$ if the good is allocated. Therefore,

$$\int_{[0,1]^n} \sum_{i=1}^n x_i q_i(x_i, x_{-i}|G) \prod_{i=1}^n (dG(x_i)) \leq \int_0^1 x dG^n \leq \int_0^1 x d\hat{G}^n = \int_{[0,1]^n} \sum_{i=1}^n x_i q_i(x_i, x_{-i}|G) \prod_{i=1}^n (d\hat{G}(x_i)).$$

The first inequality follows because the good may not be allocated under q_i . The second inequality follows because $\hat{G}$ is a mean-preserving spread of G , and because $\int_0^1 x dG^n = \int_0^1 \max\{x_1, \dots, x_n\} dG$ has an integrand that is convex in x . Moreover, the second inequality is strict if $\hat{G}$ is a strict mean-preserving spread of G . The equality follows because $\hat{G} \in \mathcal{G}_H^+$, and the good is always sold under q_i except when $x_i = 0$ for every i . (In this case, the total surplus remains the same, whether the good is sold or not.) $\square$

3 Main results

In this section, we present our results on both information design problems. We will detail their proofs in Section 4.

The unique symmetric seller-worst information structure is a truncated Pareto distribution. The distribution is regular and induces only two virtual values, k_s and 1, on the support $[x_s, 1]$, where (k_s, x_s) depends on the prior mean p and the number of buyers n (but we will omit the dependence to simplify the notations). The following result summarizes the seller-worst information structure:

Theorem 1. *For each p and n , there exists a tuple $(k_s, x_s) \in [0, 1]^2$ such that the unique symmetric seller-worst information structure is the following truncated Pareto distribution:*

$$G_s(x) = \begin{cases} 1 - \frac{x_s - k_s}{x - k_s}, & \text{if } x \in [x_s, 1]; \\ 1, & \text{if } x = 1. \end{cases} \quad (3)$$

Moreover, there exists a threshold p_s such that (i) $k_s = 0$ if $p \in (0, p_s]$; and (ii) $k_s > 0$ if $p \in (p_s, 1)$.

⁸We say that $\hat{G}$ is a mean-preserving spread of signal distribution G if $\int_0^x (\hat{G}(t) - G(t))dt \geq 0$ for all x with equality at $x = 1$; we also say that $\hat{G}$ is a strict mean-preserving spread if $\int_0^x (\hat{G}(t) - G(t))dt \geq 0$ for all x with strict inequality at some x with G -positive probability.

Theorem 1 states that the symmetric seller-worst signal distribution is a truncated Pareto distribution which induces virtual value $x - \frac{1-G_s(x)}{g_s(x)} = x - \frac{x_s/(x-k_s)}{x_s/(x-k_s)^2} = k_s$ for $x \in [x_s, 1)$, and virtual value 1 for $x = 1$. Let $\theta_s \equiv 1 - \frac{x_s - k_s}{1 - k_s}$ be the probability of G_s inducing the low virtual value k_s . Since the seller-worst information induces only nonnegative values, the good is always sold. Moreover, since the seller-worst information is regular, the seller's optimal revenue can be achieved by a second-price auction with no reserve; see Proposition 5.2 of Krishna (2009). The seller earns the expected virtual surplus equal to:

$$k_s \times \theta_s^n + 1 \times (1 - \theta_s^n). \quad (4)$$

Since G_s is the seller-worst information structure, the tuple (k_s, x_s) or, equivalently, (k_s, θ_s) is chosen to minimize the seller's expected virtual surplus, subject to the mean constraint

$$(1 - k_s)(1 - \theta_s)(1 - \log(1 - \theta_s)) + k_s = p. \quad (5)$$

The Lagrangian is linear in the low virtual value with the marginal effect being determined by the following term:⁹

$$J_s(\theta_s) \equiv \underbrace{\theta_s \log(1 - \theta_s)}_{\text{cost}(<0)} + \underbrace{n(\theta_s + (1 - \theta_s) \log(1 - \theta_s))}_{\text{benefit}(>0)}. \quad (6)$$

Intuitively, raising the low virtual value k_s has two countervailing effects on a revenue-minimizing information designer. First, by increasing the lower virtual value, the seller's revenue increases, which translates into a cost in proportion to the first term in (6). Second, to obey the mean constraint, the probability of the high virtual value 1 must be reduced, which results in a benefit in proportion to the second term in (6). Observe that the cost and benefit only depend on the number of buyers n and probability θ_s .

As the Lagrangian is linear in the low virtual value, an interior solution occurs only when the benefit exactly offsets the cost, namely $J_s(\theta_s) = 0$, and then k_s is pinned down by the mean constraint. With no more than two buyers, we can verify that the cost dominates the benefits regardless of θ_s ; hence, we must have a corner solution $k_s = 0$ and set $p_s = 1$. Consequently, the seller-worst information structure is the same for $n = 1$ and $n = 2$. If there are more than three buyers, by setting $k_s = 0$ in Equation (5), we obtain the threshold,

$$p_s = (1 - \theta_s)(1 - \log(1 - \theta_s)), \quad (7)$$

where θ_s is obtained from $J_s(\theta_s) = 0$. Again, for any prior mean below the threshold, the cost dominates the benefit and thereby we have a corner solution. We provide details in Appendix A.7.

The cost-benefit analysis also tells us how the seller-worst information varies with the prior mean or the number of buyers. The cost-benefit formula $J_s(\theta_s)$ only depends on n , and so does the probability θ_s of the low virtual value. Hence, for any fixed number of buyers, k_s must go up in proportion with p to obey the mean constraint (5). If n goes up, the benefit in $J_s(\theta_s)$ increases; hence, the probability of the low virtual value should also be increased to rebalance the cost and the benefit. Consequently, k_s

⁹To be clear, the Lagrangian is defined to maximize the additive inverse of the objective in (4).

should be raised to obey the mean constraint (5). In summary, the low virtual value k_s is increasing in both p and n .

We now turn to the buyer-optimal information. The unique buyer-optimal information structure is also a truncated Pareto distribution; however, it may place some mass θ_0 at signal $x = 0$. The distribution is regular except at $x = 0$ and induces only two virtual values k_b and 1 on the support $[x_b, 1]$, where (θ_0, k_b, x_b) depends on p and n . The following result summarizes the buyer-optimal information structure:

Theorem 2. *For each p and n , there exists a tuple $(\theta_0, k_b, x_b) \in [0, 1]^3$ such that the unique symmetric buyer-optimal information structure is the following truncated Pareto distribution (except at $x = 0$):*

$$G_b(x) = \begin{cases} \theta_0, & \text{if } x \in [0, x_b); \\ 1 - \frac{(x_b - k_b)(1 - \theta_0)}{x - k_b}, & \text{if } x \in [x_b, 1); \\ 1, & \text{if } x = 1. \end{cases} \quad (8)$$

Moreover, there exist two thresholds r_b and p_b such that (i) $k_b = 0$ and $\theta_0 > 0$ if $p \in (0, r_b)$; (ii) $k_b = 0$ and $\theta_0 = 0$ if $p \in [r_b, p_b]$; and (iii) $k_b > 0$ and $\theta_0 = 0$ if $p \in (p_b, 1)$.

The buyer-optimal signal distribution places mass θ_0 at $x = 0$ and then becomes a truncated Pareto distribution which induces virtual value k_b for $x \in [x_b, 1)$, and virtual value 1 for $x = 1$. Under the buyer-optimal information, the seller's optimal revenue can be achieved by a second-price auction with a positive reserve price $x_b > 0$; see Proposition 5.2 of Krishna (2009).

To maximize the buyers' surplus, the information designer must consider not only the seller's revenue but also the expected total surplus. As a result, the buyer-optimal information structure differs from the seller-worst information structure in several ways. First, when the prior mean $p < r_b$, the buyer-optimal signal distribution puts a mass θ_0 on signal 0; that is, with probability θ_0^n the good is not sold. Second, when the prior mean $p > p_b$, the low virtual value becomes positive; this reflects the same intuition as that of the seller-worst information structure when $p > p_s$.

Given G_b , we can compute the buyers' total surplus:

$$n(1 - k_b)(1 - \theta_b) \left(\sum_{i=1}^{n-1} \frac{-\theta_b^i}{i} - \log(1 - \theta_b) + \left(\sum_{i=1}^{n-1} \frac{\theta_0^i}{i} + \log(1 - \theta_0) \right) \right),$$

where $\theta_b \equiv 1 - \frac{(1 - \theta_0)(x_b - k_b)}{1 - k_b}$ is the mass θ_0 plus the probability of G_b inducing virtual value k_b . If $\theta_0 = 0$, then the trade-off between θ_b and k_b can be analyzed similarly as that of θ_s and k_s in the seller-worst information structure. In particular, an interior solution (k_b, θ_b) is jointly determined by the mean constraint

$$(1 - k_b)(1 - \theta_b)(1 - \log(1 - \theta_b)) + k_b = p, \quad (9)$$

and the cost-benefit equation

$$\underbrace{\theta_b \log(1 - \theta_b)}_{\text{cost}(<0)} + \underbrace{\theta_b^{n-1}(\theta_b + (1 - \theta_b) \log(1 - \theta_b)) + \sum_{i=1}^{n-1} \frac{\theta_b^{i+1}}{i}}_{\text{benefit}(>0)} = 0. \quad (10)$$

Again, since Lagrangian is linear in k_b , Equation (10) balances the cost and benefit. With no more than two buyers, we can again verify that the cost dominates the benefits regardless of θ_b ; hence, we must have a corner solution $k_b = 0$ and set $p_b = 1$.¹⁰ If there are three or more buyers, by setting $k_b = 0$ in Equation (9), we obtain the following threshold:

$$p_b = (1 - \theta_b)(1 - \log(1 - \theta_b)), \quad (11)$$

where θ_b solves Equation (10). Then, for any prior mean below the threshold p_b , the cost dominates the benefit and thereby we have a corner solution for k_b .

The cost-benefit equations in (6) and (10) also reveal that the buyer-optimal information designer is more reluctant to raise the low virtual value than a seller-worst information designer. Indeed, controlling the marginal cost, the buyer-optimal information designer receives less marginal benefit from raising the low virtual value than a seller-worst information designer does; see Claim 1 in Appendix A.8 for a formal comparison. This is because raising the low virtual value, subject to the mean constraint, results in a mean-preserving contraction; hence, it decreases the expected total surplus by Lemma 2. It also follows from the benefit comparison that p_b is larger than p_s , and for any given p , we have $k_b \leq k_s$.

As in the seller-worst information design, the cost-benefit Equation (10) only depends on n , and so does the probability θ_b . Hence, for any fixed number of buyers, k_b must go up in proportion with p to obey the mean constraint (9). If n grows up, the benefit in (10) increases; hence, the probability of the low virtual value should also be increased to rebalance the cost and the benefit. Consequently, k_b should be raised to obey the mean constraint (9). In summary, the low virtual value k_b is increasing in both p and n .

Moreover, we identify another threshold $r_b < p_b$ below which the buyer-optimal information designer also deviates from the unique seller-worst information structure. In particular, for $p < r_b$, the buyer-optimal information designer puts a positive mass θ_0 on $x = 0$ to induce a mean-preserving spread from the seller-worst information structure. While the mean-preserving spread generates more revenue for the seller, it generates even more expected total surplus to benefit the buyers. We provide a similar cost-benefit analysis to pin down (θ_0, θ_b) as well as r_b in Appendix A.8.¹¹

As the number of buyer goes to infinity, both θ_s (which solves $J_s(\theta_s) = 0$), and θ_b (which solves Equation (10)) converge to 1; hence, both p_s in (7) and p_b in (11) converge to zero. It then follows from Equations (5) and (9) that k_s and k_b will also increase to p . We summarize this in Corollary 1.

Corollary 1. *As $n \rightarrow \infty$, the buyer-optimal information structure coincides with the seller-worst information structure in the limit. Both are given by the degenerate distribution which assigns probability one to p .*

¹⁰In contrast to the seller-worst case, however, the buyer-optimal information structure for $n = 2$ may differ from that of $n = 1$, since the former may put a mass at $x = 0$.

¹¹We also show in Appendix A.8 that θ_0 and k_b cannot both be positive and thereby each cost-benefit tradeoff involves only two of the three parameters.

Corollary 1 implies that when n is large, both the buyer-optimal and the seller-worst information structures are close to “no disclosure”; that is, the information designer chooses the degenerate distribution which concentrates on $x = p$. In the limit, the seller charges and extracts the ex ante expectation p of a single buyer’s value and leaves no surplus to the buyers. This is consistent with Part (ii) of Theorem 5 in [Ganuza and Penalva \(2010\)](#) which establishes that in a second-price auction with no reserve and a sufficiently large number of buyers, a less precise signal produces a lower revenue for the seller.

Corollary 1 also contrasts with the result of [Yang \(2018\)](#). Specifically, [Yang \(2018\)](#) shows that when buyers engage in strategic information acquisition, the unique symmetric equilibrium information structure converges to full information, as the number of buyers goes to infinity. Hence, the buyers also retain zero surplus in the limit, as in our buyer-optimal information structure. Since our symmetric buyer-optimal information structure provides an upper bound of the buyers’ surplus under the symmetric equilibrium in [Yang \(2018\)](#), the comparison reveals that their gap vanishes as the number of buyers goes to infinity.¹²

In [Yang \(2018\)](#), the buyers’ surplus goes to zero because of the increasing competition in information acquisition with more buyers. In our case, however, the limiting zero surplus is driven by the buyer-optimal information designer’s intentional choice to increase the low virtual value k_b in order to reduce the probability of virtual value 1. As we explained after presenting Theorem 1, when n is large, reducing the probability of a high virtual value becomes a dominant consideration for the buyer-optimal information designer. Nevertheless, we will show later that the buyers do retain a nonvanishing surplus even when n goes to infinity; as long as the information designer can commit herself to choosing an asymmetric information structure; see Section 5.2.

Figure 1 provides a numerical example with $p = 1/2$ and $n = 1, 2, \dots, 10$. We document a number of features in this example to illustrate our main results.

- The top subfigure shows:
 - As n increases, the low virtual value goes up in both the seller-worst and buyer-optimal cases.
 - The seller-worst low virtual value k_s is larger than the buyer-optimal low virtual value k_b .
 - k_s tends to 0.5 faster than k_b .
 - When $n = 2$, the buyer-optimal signal distribution puts a mass of 0.06 on $x = 0$ against which the seller withholds the good.
- The bottom-left subfigure shows:
 - The buyers’ surplus under the buyer-optimal information structure is strictly higher than that of the seller-worst information structure.

¹²[Shi \(2012\)](#) studies an optimal auction problem in which the buyers acquire information individually from restricted feasible information structures *after* the auction is announced; in contrast, in [Yang \(2018\)](#) as well as in our paper, the optimal auction is designed according to the information, whether by design in our case or from the buyers’ strategic acquisition in the case of [Yang \(2018\)](#).

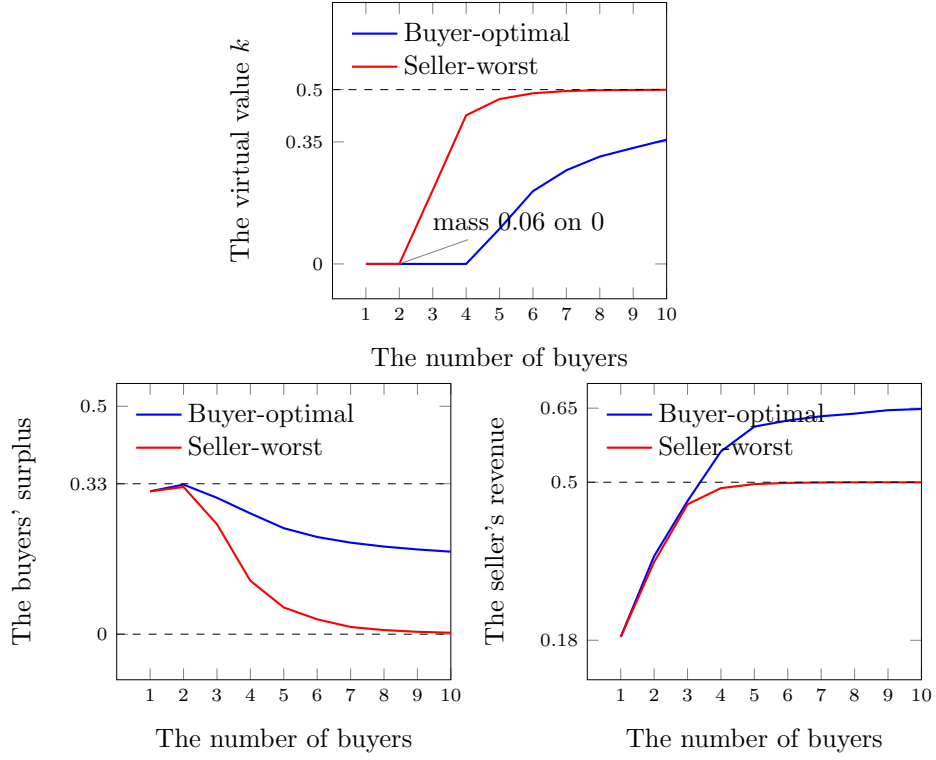


Figure 1: A simulation for different n with $p = 0.5$

- The buyers' surplus grows only from $n = 1$ to $n = 2$ and then decreases with n .
- The bottom-right subfigure shows:
 - The seller's revenue under the buyer-optimal information structure is strictly higher than that of the seller-worst information structure.
 - The seller-worst revenue never exceeds 0.5 since the information designer can opt to disclose no information. In contrast, with $n \geq 4$, the seller's revenue under the buyer-optimal information exceeds 0.5.

4 Outline of the solution

Here we outline the steps to solve the information design problems we encounter. As we mentioned in the introduction, the key idea is to reduce the problems into tractable, finite-dimensional constrained optimization problems.

1. We first present two preliminary lemmas to restrict the class of distributions of interest:
 - (a) Any buyer-optimal information structure induces nonnegative virtual values except at 0, i.e., the buyer-optimal signal distribution must belong to $\mathcal{G}_H^+$. Moreover, any seller-worst information structure must induce nonnegative virtual values almost everywhere. (Lemma 4).

- (b) Any buyer-optimal signal distribution must be “regular except at 0” (i.e., regular everywhere except at $x = 0$) and any seller-worst signal distribution must be regular (Lemma 5).
- 2. We change the choice variable in our information design problems from the distribution of signals to the distribution of virtual values (Lemma 7).
- 3. We show that the reformulation after the change of variable reduces the information design problems to tractable, finite-dimensional optimization problems. This enables us to derive the explicit solution to the information design problems, regardless of the number of buyers. In particular, when $n = 1$, we obtain the same solution that appears in Roesler and Szentes (2017).

4.1 Preliminary Lemmas

In this subsection, we present the two preliminary results, Lemmas 4 and 5, to restrict the class of distributions of interest to the information designer. Both results rely on the following observation established by Roesler and Szentes (2017) and Yang (2018).

Denote the left- and right-hand limit of a function $\xi(\cdot)$ at a signal x by $\xi(x^-) = \lim_{\delta \downarrow 0} \xi(x - \delta)$ and $\xi(x^+) = \lim_{\delta \downarrow 0} \xi(x + \delta)$, respectively.

Lemma 3. *For any distribution G , any $x_0 \in [0, 1]$, and any $k \in [\hat{\varphi}(x_0^-|G), \hat{\varphi}(x_0^+|G)]$,*

$$(x - k)(1 - G(x)) \leq (x_0 - k)(1 - G(x_0^-)), \quad \forall x \in [0, 1]. \quad (12)$$

Proof. See Lemma 9 in Yang (2018). □

We can rearrange the inequality (12) to obtain

$$G(x) \geq 1 - \frac{(x_0 - k)(1 - G(x_0^-))}{x - k}.$$

In particular, the right-hand side is a Pareto distribution function which generates a constant virtual value k . Hence, Lemma 3 says that the Pareto distribution first-order stochastically dominates any other distribution G with $k \in [\hat{\varphi}(x_0^-|G), \hat{\varphi}(x_0^+|G)]$.

Lemma 4. *Any optimal signal distribution G which solves the information design problem in (1) must induce nonnegative virtual values except at 0 (i.e., $G \in \mathcal{G}_H^+$); moreover, if G is a solution to (1) for $\alpha = 0$, it must induce nonnegative virtual values almost everywhere on $[0, 1]$.*

Proof. See Appendix A.3. □

Figure 2 illustrates that how to improve the information designer’s objective (for both the buyer-optimal case and the seller-worst case). First, in the left subfigure, the blue curve is an arbitrary candidate distribution, which generates negative virtual values over the interval $[0, x_2]$. The red curve coincides with the blue curve for $x \geq x_2$. For $x < x_2$, the red curve is a Pareto distribution which generates a virtual value 0 over the interval (x_1, x_2) . By Lemma 3, we also put a positive mass θ_0 on

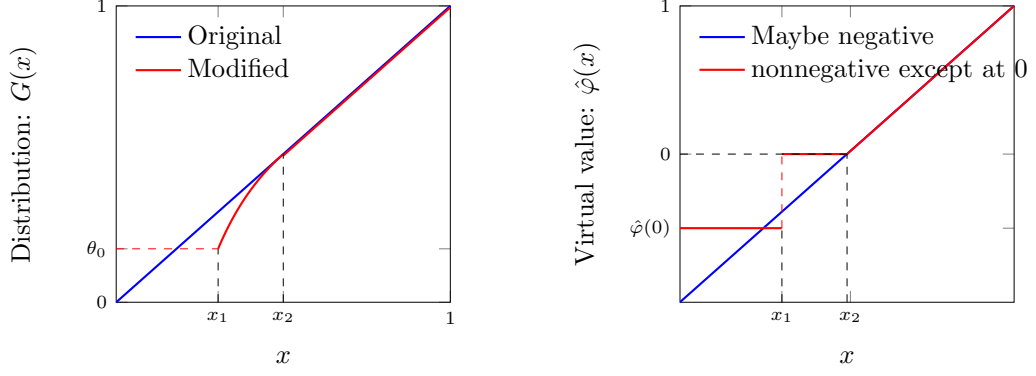


Figure 2: An improvement by a distribution with nonnegative virtual value except at 0

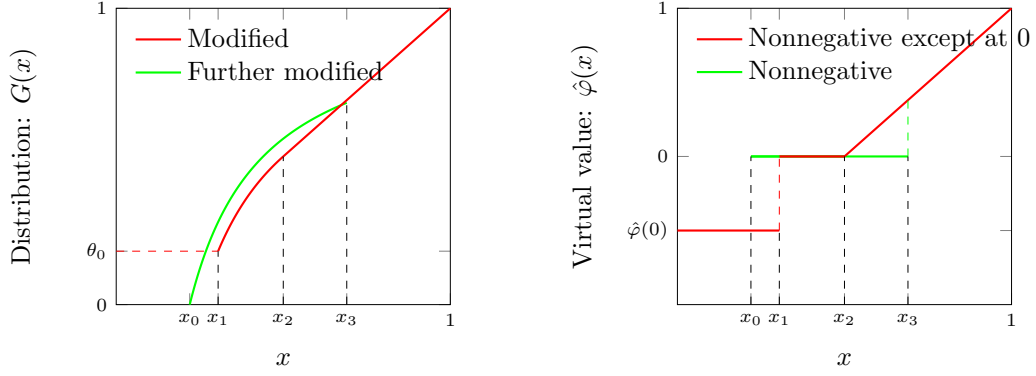


Figure 3: A further improvement by a nonnegative virtual value distribution

$x = 0$ for the red curve to generate the same mean as the blue curve. By construction, the red curve is a strict mean-preserving spread of the blue curve; therefore, by Lemma 2, the red curve can generate more expected total surplus. Second, by construction, for $x \in (0, x_2)$, both of the ironed virtual values are no more than zero, and for $x \in [x_2, 1]$ the nonnegative ironed virtual value of the red curve coincides with that of the blue curve in the right subfigure. Therefore, the red curve generates strictly higher surplus for the buyers than the blue curve. In summary, the modification strictly benefits a buyer-optimal information designer and causes no loss to a seller-worst information designer. Hence, a buyer-optimal information structure must induce nonnegative virtual values except at 0.

Even though the seller-worst information designer does not benefit from the modification in Figure 2, we use Figure 3 to illustrate how we can further modify the red curve in Figure 2 to decrease the seller's revenue. Intuitively, since the seller gets no revenue as long as the virtual value is nonpositive, instead of inducing a negative virtual value at zero, a seller-worst information designer can do better by raising the negative virtual value to zero while decreasing the probability assigned to a positive virtual value to maintain the mean. More precisely, consider the green curve, which coincides with the red curve for $x \geq x_3$. For $x < x_3$, the green curve becomes a Pareto distribution which generates virtual value 0 over the interval (x_0, x_3) such that the green curve generates the same mean as the red curve. By

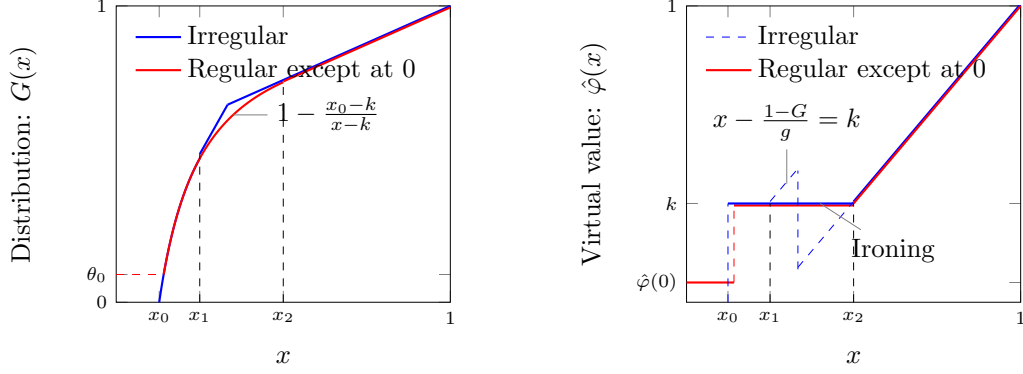


Figure 4: An improvement by a regular distribution except at 0

construction, and as shown in the right subfigure, the green curve generates strictly lower virtual values than the red curve over the interval (x_2, x_3) ; hence, the green curve generates strictly less revenue.

Lemma 5. *Any optimal signal distribution G which solves the information design problem in (1) must be regular except at 0; moreover, if G is a solution to (1) for $\alpha = 0$, it must be regular.*

Proof. See Appendix A.4. □

Figure 4 illustrates how to improve the buyer-optimal information designer's objective with regular distributions except at 0. First, in the left subfigure, the blue curve is a candidate distribution, which is irregular over the interval (x_1, x_2) and generates ironed virtual value $k \geq 0$ over the interval $[x_0, x_2]$. The red curve coincides with the blue curve for $x \geq x_2$. For $x < x_2$, the red curve becomes a Pareto (and hence regular) distribution which generates a same virtual value k over the interval (x_1, x_2) and puts remaining mass θ_0 on 0. Observe that the red curve is a strict mean-preserving spread of the blue curve in the left subfigure.¹³ Hence, by Lemma 2, the red curve can generate more expected total surplus. Second, the red curve induces weakly lower ironed virtual values than the blue curve; hence, the seller earns less under the red curve than under the blue curve. Overall, the buyer-optimal information designer's objective value is higher under the red curve than under the blue curve. Moreover, by Lemma 4, the red curve induces negative virtual value at $x = 0$ and hence neither the red curve nor the blue curve can be seller-worst.

By Lemma 5, we will hereafter use φ instead of $\hat{\varphi}$ to denote the virtual value.

4.2 Change of variable

We now introduce a key step in solving the information design problems, namely, we change our control variable from a signal distribution to a virtual value distribution. Let $F(k)$ be the distribution of virtual

¹³We draw the blue curve above the red curve on the ironed interval $[x_1, x_2]$ because the Pareto signal distribution with virtual value k first-order stochastically dominates the original signal distribution by Lemma 3.

values given a feasible signal distribution G . Since G is regular except at 0, $F(k) = \text{Prob}_G \{x | \varphi(x) \leq k\}$. Then, except at $x = 0$, the virtual value of G at signal x is

$$\varphi(x) = x - \frac{1 - G(x)}{G'(x)}. \quad (13)$$

First, assume that the virtual valuation function $\varphi(\cdot)$ is *strictly increasing*, so that its inverse $x(k) = \varphi^{-1}(k)$ is well defined. Consequently, $F(k) = G(x(k))$ and $F'(k) = G'(x(k))x'(k)$. By Equation (13), we have the following ordinary differential equation of $x(k)$:

$$k = x(k) - \frac{1 - F(k)}{F'(k)} x'(k). \quad (14)$$

Solving the differential equation, we obtain Lemma 6 below.

Lemma 6. *Suppose that φ is strictly increasing in x . For each k with $\varphi(x) = k$, $x(k)$ is the buyer's expected virtual value conditional on his virtual value being greater than or equal to k , i.e.,*

$$x(k) = \mathbb{E}[\varphi | \varphi \geq k] = k + \frac{\int_k^1 (1 - F(s)) ds}{1 - F(k)}. \quad (15)$$

Alternatively, we can also derive the expression of $x(k)$ by applying the Envelope Theorem.¹⁴ Observe that $x(k)$ is the solution to the following monopoly pricing problem with marginal cost k :

$$V(k) = \max_x (x - k)(1 - G(x)).$$

By the Envelope Theorem, we can derive:

$$\frac{\partial V(k)}{\partial k} = -(1 - G(x(k))).$$

Hence,

$$\begin{aligned} V(1) - V(k) &= -(x(k) - k)(1 - G(x(k))) = \int_k^1 -(1 - G(x(s))) ds \\ \implies x(k) &= k + \frac{\int_k^1 (1 - G(x(s))) ds}{1 - G(x(k))} = k + \frac{\int_k^1 (1 - F(s)) ds}{1 - F(k)}. \end{aligned}$$

Moreover, we can still make use of the expression for $x(k)$ in (15), even when $\varphi(x)$ is only weakly increasing (by regularity). To see this, note that any weakly increasing function can be uniformly approximated by a strictly increasing function. Let $\{\varphi_m\}_{m=1}^\infty$ be a sequence of strictly increasing functions converging uniformly to φ . For each m , let G_m and F_m be sequences of signal distributions and virtual value distributions corresponding to φ_m . Specifically, by solving Equation (13), we obtain $G_m(x) = 1 - \exp \left\{ \int_0^x (\varphi_m(t) - t)^{-1} dt \right\}$ and F_m is the virtual value distribution induced by G_m . Since $\{\varphi_m\}$ converges uniformly to φ , we also have $\{G_m\}$ and $\{F_m\}$ uniformly converge to G and F , respectively.

In Appendix A.5, we use this convergence result and the expression in (15) to establish the following

¹⁴We thank Kai Hao Yang for suggesting this elegant argument to us.

equation:

$$\begin{aligned} \int_0^1 x dG^n(x) &= 1 - \int_0^1 G^n(x) dx \\ &= \int_0^1 n(1 - F(k)) \left(\sum_{i=1}^{n-1} \frac{-F^i(k)}{i} - \log(1 - F(k)) + \sum_{i=1}^{n-1} \frac{F^i(0^-)}{i} + \log(1 - F(0^-)) \right) - F^n(k) dk + 1, \end{aligned} \quad (16)$$

where $G(0) = F(0^-)$. Indeed, we show in Appendix A.5 that Equation (16) follows for each G_m and F_m and we then apply the bounded convergence theorem. To see the intuition, suppose that $n = 1$ and there is no mass on signal 0. Then, Equation (16) is reduced to:

$$\int_0^1 (1 - F(k))(1 - \log(1 - F(k))) dk = p. \quad (17)$$

We may regard (17) as a mean constraint for a general virtual value distribution that generalizes the special case (5) for a binary virtual value distribution on $\{k_s, 1\}$.

Equation (16) is the total surplus given the distribution G , and when $n = 1$, Equation (16) is reduced to the expected mean. Hence, we have the following lemma.

Lemma 7. *After the change of variable, the information designer's problem in (1) can be written as follows:*

$$\begin{aligned} \max_F \int_0^1 \alpha n(1 - F(k)) &\left(\sum_{i=1}^{n-1} \frac{-F^i(k)}{i} - \log(1 - F(k)) + \left(\sum_{i=1}^{n-1} \frac{F^i(0^-)}{i} + \log(1 - F(0^-)) \right) \right) dk \\ &+ (1 - \alpha) \int_0^1 F^n(k) dk + (\alpha - 1) \\ \text{s.t. } \int_0^1 &(1 - F(k))(1 - \log(1 - F(k)) + \log(1 - F(0^-))) dk = p. \end{aligned} \quad (18)$$

Proof. It directly follows from Equation (16). $\square$

4.3 The case with $n = 1$: Roesler and Szentes (2017) revisited

We are now ready to solve the information design problem for the case with $n = 1$, which is analyzed in Roesler and Szentes (2017). For $n = 1$, we can rewrite the information designer's problem as

$$\begin{aligned} \max_F \quad & \alpha \underbrace{\int_0^1 x dG(x)}_{\text{the total surplus}} - \underbrace{\int_0^1 k dF(k)}_{\text{the seller's revenue}} \\ \text{s.t. } \quad & \int_0^1 (1 - F(k))(1 - \log(1 - F(k)) + \log(1 - F(0^-))) dk = p. \end{aligned}$$

When $n = 1$, the total surplus $\int_0^1 x dG$ is linear in G ; moreover, $\int_0^1 x dG = p$ by the mean constraint. Therefore, the value of α has no effect on the optimization. This implies that the buyer-optimal information structure is equivalent to the seller-worst information structure. Since the virtual value is always nonnegative for the seller-worst case, we have $G(0) = F(0^-) = 0$, i.e., there is no mass on $x = 0$.

The information design problem is an isoperimetric problem in optimal control theory; see Theorem 4.2.1 of [van Brunt \(2004\)](#). To solve the optimal control problem, we can write the Lagrangian as

$$\mathcal{L}(F, \lambda) = \alpha p + \int_0^1 (F - \lambda((1 - F)(1 - \log(1 - F)) - p)) dk - 1.$$

Let $\theta = F(k)$. Then, for each k , the Euler-Lagrange condition implies that

$$\partial \mathcal{L} / \partial \theta = 1 - \lambda \log(1 - \theta) = 0.$$

Since the solution of the Euler-Lagrange equation cannot be either $\theta = 0$ or $\theta = 1$, there exists a constant $\lambda = 1 / \log(1 - \theta) < 0$. Since $\log(1 - \theta)$ is monotone decreasing, there is at most a solution $\theta \in (0, 1)$, such that the Euler-Lagrange equation holds. Therefore, $F(k)$ has only three values 0, θ , and 1; hence, F is a two-point distribution. Moreover, since the constant λ is fixed, and since $\log(1 - \theta)|_{\theta \uparrow 1} \rightarrow -\infty$, we have

$$\left. \frac{\partial \mathcal{L}}{\partial \theta} \right|_{\theta \uparrow 1} = 1 - \lambda \log(1 - \theta)|_{\theta \uparrow 1} \rightarrow -\infty.$$

Therefore, it will never be optimal for $F(k)$ to jump to 1 before $k = 1$; therefore, the larger value with respect to F must be 1. That is, the support of F is $\{k, 1\}$.

Now, the information designer only needs to choose $\{k, F(k) = \theta\}$ to maximize

$$\begin{aligned} \max_{k \geq 0, \theta} \quad & \alpha p - (\theta \times k + (1 - \theta) \times 1) \\ \text{s.t.} \quad & k + (1 - k)(1 - \theta)(1 - \log(1 - \theta)) = p. \end{aligned}$$

The Lagrangian is

$$\mathcal{L}(k, \theta, \lambda, \mu) = \alpha p - k\theta - (1 - \theta) + \lambda(p - (k + (1 - k)(1 - \theta)(1 - \log(1 - \theta)))).$$

and the Euler-Lagrange equation for θ is

$$\frac{\partial \mathcal{L}}{\partial \theta} = (1 - k)(1 - \lambda(\log(1 - \theta))) = 0.$$

Therefore,

$$\lambda = 1 / \log(1 - \theta).$$

Hence, the Euler-Lagrange equation for k is

$$\frac{\partial \mathcal{L}}{\partial k} = -\theta - \lambda(\theta + (1 - \theta)\log(1 - \theta)) = \frac{\theta + \log(1 - \theta)}{-\log(1 - \theta)} < 0.$$

Therefore, the optimal $k = 0$, and F is a two-point distribution which puts mass only on the virtual values 0 and 1.

In summary, we have reproduced the optimal signal distribution derived in [Roesler and Szentes \(2017\)](#), namely,

$$G(x) = \begin{cases} 1 - \frac{1-\theta}{x}, & \text{if } x \in [1 - \theta, 1); \\ 1, & \text{if } x = 1. \end{cases}$$

Under the optimal signal distribution, for $x \in [1 - \theta, 1)$, the virtual value is zero with probability θ . For $x = 1$, the virtual value is 1 with probability $1 - \theta$.

4.4 The case with $n \geq 2$

As in the case with $n = 1$, we can reduce the infinite-dimensional information design problem to a finite-dimensional problem by the following lemma.

Lemma 8. *The support of any optimal virtual value distribution F has two points, say, $\{k, 1\}$.*

Proof. The information design problem is also an isoperimetric problem in optimal control theory. Define the following Lagrangian,

$$\begin{aligned} \mathcal{L}(F, \lambda) = & \int_0^1 \alpha n(1 - F(k)) \left(\sum_{i=1}^{n-1} \frac{-F^i(k)}{i} - \log(1 - F(k)) + \left(\sum_{i=1}^{n-1} \frac{F^i(0^-)}{i} + \log(1 - F(0^-)) \right) \right) dk \\ & + \int_0^1 (1 - \alpha) F^n(k) - \lambda(1 - F(k))(1 - \log(1 - F(k)) + \log(1 - F(0^-))) dk + p\lambda + (\alpha - 1). \end{aligned}$$

Let $\theta_0 = F(0^-) = G(0)$, and let $\theta = F(k)$. By Theorem 4.2.1 of [van Brunt \(2004\)](#), the Euler-Lagrange equation for $\mathcal{L}$ and for each state k should be satisfied as follows:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \theta} = & -n\alpha \left(\sum_{i=1}^{n-1} -\theta^i/i - \log(1 - \theta) + \left(\sum_{i=1}^{n-1} \theta_0^i/i + \log(1 - \theta_0) \right) \right) + n\alpha(1 - \theta) \left(\sum_{i=1}^{n-1} -\theta^{i-1} + \frac{1}{1 - \theta} \right) \\ & + n(1 - \alpha)\theta^{n-1} - \lambda \log(1 - \theta) + \lambda \log(1 - \theta_0) \\ = & n\alpha \sum_{i=1}^{n-1} \theta^i/i + n\alpha \log(1 - \theta) - \lambda \log(1 - \theta) + \left(-n\alpha \sum_{i=1}^{n-1} \theta_0^i/i - (n\alpha - \lambda) \log(1 - \theta_0) \right) \\ & + n\alpha(1 - \theta) \left(\frac{\theta^{n-1} - 1}{1 - \theta} + \frac{1}{1 - \theta} \right) + n(1 - \alpha)\theta^{n-1} \\ = & n\alpha \sum_{i=1}^{n-1} \theta^i/i + n\theta^{n-1} + n\alpha \log(1 - \theta) - \lambda \log(1 - \theta) + \left(-n\alpha \sum_{i=1}^{n-1} \theta_0^i/i - (n\alpha - \lambda) \log(1 - \theta_0) \right) = 0. \end{aligned}$$

Denote $\partial \mathcal{L} / \partial \theta$ by $I_\alpha(\theta)$. Then by taking the derivative of $I_\alpha(\theta)$ with respect to θ , we have:

$$\begin{aligned} I'_\alpha(\theta) = & n\alpha \sum_{i=1}^{n-1} \theta^{i-1} + n(n-1)\theta^{n-2} + \frac{-n\alpha + \lambda}{1 - \theta} = \frac{n\alpha(1 - \theta^{n-1})}{1 - \theta} + n(n-1)\theta^{n-2} + \frac{-n\alpha + \lambda}{1 - \theta} \\ = & \frac{-n\alpha\theta^{n-1} + n(n-1)\theta^{n-2}(1 - \theta) + \lambda}{1 - \theta} = \frac{\lambda + n\theta^{n-2}(n-1 - (n-1 + \alpha)\theta)}{1 - \theta}. \end{aligned}$$

We prove the following lemma in [Appendix A.6](#):

Lemma 9. *There is at most one θ with $I_\alpha(\theta) = 0$ which also satisfies the second-order condition.*¹⁵

Therefore, $F(k)$ has only three values 0, θ , and 1; hence, F is a two-point distribution. We then argue that for both $\alpha = 0$ and $\alpha = 1$, signal 1 is on the support of F .

1. When $\alpha = 0$, it is only when $\lambda < 0$ that feasible solutions of θ exist; see cases 2 and 3 in [Appendix A.6](#). Therefore,

$$\begin{aligned} I_0(\theta)|_{\theta \uparrow 1} = & n\theta^{n-1}|_{\theta \uparrow 1} - \lambda(\log(1 - \theta)|_{\theta \uparrow 1} + \lambda \log(1 - \theta_0)) \\ = & n + \lambda \log(1 - \theta_0) - \lambda \log(1 - \theta)|_{\theta \uparrow 1} \rightarrow -\infty. \end{aligned}$$

¹⁵In [Figure 5](#), we also draw the curve of $I_\alpha(\theta)$ and $I'_\alpha(\theta)$ to illustrate this lemma. In [Figure 5](#), we choose the parameters to be $n = 3$, $\alpha = 1$, $\lambda = -0.5$, and $\theta_0 = 0$.

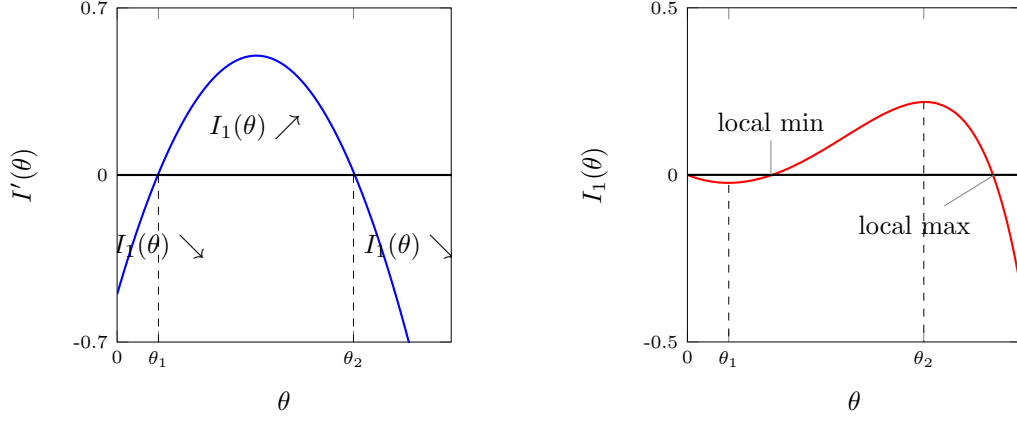


Figure 5: The curve of $I'_1(\theta)$ and $I_1(\theta)$.

Therefore, having $F(k)$ jump to 1 before $k = 1$ will never be optimal.

2. When $\alpha = 1$, it is only when $\lambda < n$, feasible solutions of θ exist; see cases 2, 3 and 4 in Appendix A.6. Therefore,

$$\begin{aligned}
I_1(\theta)|_{\theta \uparrow 1} &= \left(n \sum_{i=1}^{n-1} \theta^i / i + n\theta^{n-1} + n \log(1 - \theta) - \lambda \log(1 - \theta) \right) \Big|_{\theta \uparrow 1} \\
&\quad + \left(-n \sum_{i=1}^{n-1} \theta_0^i / i - (n - \lambda) \log(1 - \theta_0) \right) \\
&= n + n \sum_{i=1}^{n-1} 1/i + \left(-n \sum_{i=1}^{n-1} \theta_0^i / i - (n - \lambda) \log(1 - \theta_0) \right) \\
&\quad + (n - \lambda) \log(1 - \theta)|_{\theta \uparrow 1} \rightarrow -\infty.
\end{aligned}$$

Therefore, having $F(k)$ jump to 1 before $k = 1$ will never be optimal.

In summary, the support of F has two points $\{k, 1\}$. □

Therefore, the information designer will choose the optimal k, θ , and θ_0 (with $\theta = F(k)$ and $\theta_0 = G(0) = F(0^-)$) such that

$$\begin{aligned}
&\max_{\theta_0, k, \theta} \left(\alpha n(1 - k)(1 - \theta) \left(\sum_{i=1}^{n-1} \frac{-\theta^i}{i} - \log(1 - \theta) + \left(\sum_{i=1}^{n-1} \frac{\theta_0^i}{i} + \log(1 - \theta_0) \right) \right) \right) \\
&\quad + (\alpha - 1) + (1 - \alpha)(1 - k)\theta^n + (1 - \alpha)k\theta_0^n \tag{19} \\
&\text{s.t. } (1 - k)((1 - \theta)(1 - \log(1 - \theta) + \log(1 - \theta_0)) + k(1 - \theta_0) = p, \\
&\quad k \geq 0, \quad \theta_0 \geq 0, \quad \text{and} \quad \theta_0 \leq \theta \leq 1.
\end{aligned}$$

The solution to this finite-dimensional optimization problem is standard and we present the details in Appendices A.7 and A.8. In fact, the solution to this finite-dimensional problem is unique. Since the information design problem has at least one solution by Lemma 1, this unique solution is globally optimal; this concludes the proofs of Theorems 1 and 2.

We now wish to briefly comment on how our approach differs from that of [Roesler and Szentes \(2017\)](#). When there is only one buyer, a posted price mechanism is optimal. Hence, a signal distribution matters only in determining the optimal posted price, i.e., $\varphi^{-1}(0)$. This is how [Roesler and Szentes \(2017\)](#) are able to argue that it entails no loss of generality to focus on a class of Pareto distributions, and the two-point Pareto distribution with virtual values 0 and 1 is the buyer-optimal information structure. When there are multiple buyers, we may take the second-price auction with an optimal reserve price to be an extension of the posted price mechanism. Unlike [Roesler and Szentes \(2017\)](#), however, a second-price auction with a reserve price need not be an optimal auction against an irregular signal distribution.

To explain, while an irregular signal distribution can be ironed into a regular signal distribution, the optimal expected revenue under the irregular distribution is the same as the revenue of a second-price auction with a reserve price under the regular distribution obtained from ironing rather than the irregular distribution.¹⁶ We show in Section 6.2 that when the seller is committed to using a second-price auction with reserve price 0, then for $n = 2$, the seller-worst information structure is the binary prior (i.e., full revelation). Hence, contrary to Theorem 1, the information structure which minimizes the seller's revenue in a second-price auction is an irregular distribution.

5 Asymmetric information structures

So far we have assumed that the information designer chooses the same signal distribution across all buyers. A natural question to ask is whether the information designer can do better by choosing different signal distributions for different buyers. The short answer is “No” for the seller-worst information design problem and “Yes” for the buyer-optimal information design problem. We elaborate further below.

To allow for asymmetric signal distributions, we first reformulate the information design problems. Let $M(x|\mathbf{G}) = \{i \in N | \hat{\varphi}(x_i|G_i) \geq \max_j \{\hat{\varphi}(x_j|G_j), 0\}\}$ be the set of buyers who have the largest nonnegative virtual value for a given signal realization x , where $\mathbf{G}$ stands for $\{G_i\}_{i=1}^n$; and let $M'(x|\mathbf{G}) = \{i \in N | x_i \geq \max_j x_j, \forall j \in M(x|\mathbf{G})\}$ be the set of buyers who not only have the largest virtual value but also the largest signal among those with the highest virtual value for a given signal realization x . Then, the optimal auction allocation rule for buyer i when all buyers report their signals is given by the following:

$$q_i(x_i, x_{-i}|\mathbf{G}) = \begin{cases} \frac{1}{|M'(x|\mathbf{G})|}, & \text{if } i \in M'(x|\mathbf{G}); \\ 0, & \text{if } i \notin M'(x|\mathbf{G}). \end{cases}$$

¹⁶For the sake of completeness, we provide an example in Appendix A.12 to illustrate this point; see also the working paper version of [Monteiro and Svaiter \(2010\)](#) for a similar example. In particular, our example admits a density and nonnegative (ironed) virtual values for every signal.

We now study the following information design problem:

$$\begin{aligned} \max_{\{G_i\}_{i=1}^n} & \int_{[0,1]^n} \sum_{i=1}^n (\alpha x_i - \hat{\varphi}(x_i|G_i)) q_i(x_i, x_{-i}|\mathbf{G}) \prod_{i=1}^n (dG_i(x_i)) \\ \text{s.t.} & \int_0^1 1 - G_i(x_i) dx_i = p, \quad \forall i = 1, \dots, n. \end{aligned} \quad (20)$$

While we allow for asymmetric signal distributions in the information design problem (20), we still assume that the agents' binary priors have the same mean p . We relegate the discussion about asymmetric priors to Section 6.1.

5.1 The seller-worst case

In this section, we show that the optimal symmetric seller-worst information structure in Theorem 1 remains the unique optimal solution, even if the information designer can choose an asymmetric information structure. We first document the existence of the solution to problem (20) when $\alpha = 0$; see Lemma 10. The proof is similar to the proof of Lemma 1 and is therefore omitted.

Lemma 10. *For the problem in (20) with $\alpha = 0$, an optimal solution exists.*

The following Lemma 11, corresponds to Lemmas 4 and 5. The proof is similar so we provide only a sketch of it in Appendix A.9.

Lemma 11. *Any profile of optimal signal distributions $\{G_i\}_{i=1}^n$ which solves the information design problem in (20) must be regular and induce nonnegative virtual values almost everywhere on $[0, 1]$.*

Then, as in Lemma 7, it follows from Lemma 11 that the asymmetric information design problem can be reformulated as:

$$\begin{aligned} \max_{\{F_i\}_{i=1}^n} & \int_0^1 \prod_{i=1}^n F_i(k) dk - 1 \\ \text{s.t.} & \int_0^1 (1 - F_i(k))(1 - \log(1 - F_i(k))) dk = p, \quad \forall i = 1, \dots, n. \end{aligned} \quad (21)$$

We now state and prove the following theorem.

Theorem 3. *The unique seller-worst information structure in Theorem 1 remains the unique seller-worst information structure which solves the problem in (20) with $\alpha = 0$.*

Proof. For any profile of virtual value distributions $\{F_i\}_{i=1}^n$, let $F(k) \equiv \frac{1}{n} \sum_{i=1}^n F_i(k)$, and denote by F a symmetric signal distribution profile where each buyer receives his signal according to F .

First, the symmetric signal distribution F yields weakly less revenue than $\{F_i\}_{i=1}^n$ does. For any k , by the inequality of arithmetic and geometric means, we have

$$F^n(k) = \left(\frac{1}{n} \sum_{i=1}^n F_i(k) \right)^n \geq \prod_{i=1}^n F_i(k). \quad (22)$$

Moreover, the equality in (22) holds if and only if $F_i(k) = F$ for all i . Integrating both sides yields $\int_0^1 F^n(k)dk \geq \int_0^1 \prod_{i=1}^n F_i(k)dk$. Hence, F yields weakly less revenue than $\{F_i\}_{i=1}^n$ does and strictly less revenue when $F_i \neq F$ for some i .

Second, F generates a weakly higher mean than p . Indeed, the integral term in the mean constraint is strictly concave with respect to $F(k)$. That is, $I''(\theta) = -1/(1-\theta) < 0$ where $I(\theta) = (1-\theta)(1-\log(1-\theta))$. Next, define

$$\hat{F}(k) = \begin{cases} F(k_1^-), & \text{if } k \in [0, k_1); \\ F(k), & \text{if } k \in [k_1, 1], \end{cases}$$

for some k_1 so that $\hat{F}$ satisfies the constraint in (21). Since $\hat{F}(k) \geq F(k)$ for any k , it follows that $\int_0^1 \hat{F}^n(k)dk \geq \int_0^1 F^n(k)dk$ and $\hat{F}$ yields less revenue than F . Therefore, the improvement from $\{F_i\}_{i=1}^n$ to $\hat{F}$ is strict when $F_i \neq F$ for some i . It follows that the symmetric seller-worst information structure in Theorem 1 remains the unique solution to the problem in (21) with $\alpha = 0$. $\square$

5.2 The buyer-optimal case

For the buyer-optimal information design problem, we demonstrate that an asymmetric signal distribution can strictly improve upon the optimal symmetric signal distribution in Theorem 2. We demonstrate such an improvement for the case of $n = 2$ with $p < r_b$, and the case of $n \rightarrow \infty$.

Proposition 1. *For $n = 2$ with $p < r_b$ or for $n \rightarrow \infty$, there exist asymmetric information structures which can strictly improve upon the optimal symmetric signal distribution in Theorem 2 and Corollary 1, respectively.*

Proof. See Appendix A.10. $\square$

To see the main idea, for the case where $n = 2$ with $p < r_b$, we focus our search for improvement on the signal distributions which put a positive mass only on signal 0, on signals with virtual value 0, and on signals with virtual value 1 for each buyer. Since each buyer's virtual value distribution needs to satisfy the mean constraint, it is uniquely determined by its probability assigned to signal 0. We then optimize within the specific class of information structures with these two probabilities assigned to signal 0 (one for each buyer). It turns out that within this class of distributions, the two buyers' aggregate surplus is maximized when one buyer's signal distribution assigns positive probability on signal 0, whereas the other buyer's signal distribution assigns zero probability. Moreover, the former buyer benefits, while the latter one loses relative to the symmetric buyer-optimal information; see Appendix A.10 for more details.

For the case where $n \rightarrow \infty$, we also consider a specific class of information structures in which (i) buyer i 's signal distribution puts positive mass only on signal 0, on signals with virtual value p , and on signals with virtual value 1; and (ii) all the other buyers' signal distributions are a degenerate distribution which puts the entire mass on signal p . We will construct an asymmetric information structure in which

no buyer is worse off and some buyer is strictly better off, relative to the symmetric buyer-optimal information structure.

Proposition 1 shows that the buyer-optimal signal distributions, whether there exists one or many, are asymmetric in general. Also in general, the total surplus as a max function of buyers' value favors dispersion. Indeed, while averaging the signal distributions ($F = \frac{1}{n} F_i$) can reduce the revenue, it might reduce the total surplus. The issue is reminiscent of the result in Bergemann and Pesendorfer (2007) which shows that a seller-optimal information structure is asymmetric across all buyers. In Bergemann and Pesendorfer (2007), an information structure is chosen to maximize the expected nonnegative virtual surplus, which also favors asymmetry/dispersion. Therefore, relative to the optimal symmetric information structure, an asymmetric information structure increases the total surplus as well as the seller's revenue. Proposition 1 demonstrates that we can choose an asymmetric information structure which increases the expected total surplus more than the expected (nonnegative) virtual surplus.¹⁷

6 Discussion

So far we have studied our information design problem with ex ante symmetric binary priors and a Myersonian optimal auction. In this section, we will first discuss the issues with asymmetric priors and continuous priors. Next, we will discuss the tightness of our seller-worst revenue upper bound for the revenue guarantee of informationally robust auctions.

6.1 Asymmetric prior mean

Suppose that each buyer i has his own prior mean p_i . We discuss only the seller-worst problem for which we know a solution exists. The seller-worst information problem is to maximize the same objective function in (21) with each buyer i 's individual mean constraint now being:

$$\int_0^1 (1 - F_i(k))(1 - \log(1 - F_i(k)) + \log(1 - F_i(0^-)))dk = p_i.$$

In this case, our arguments for regularity and nonnegative virtual values are still valid except at zero.

For $n = 2$, we obtain the seller-worst information structure in the following proposition.

Proposition 2. *Suppose that $n = 2$ and each buyer i has a prior mean p_i . Then, the seller-worst information structure is a profile of signal distributions $\{G_1, G_2\}$ such that, for each buyer i , G_i is a truncated Pareto distribution as follows:*

$$G_i(x) = \begin{cases} 1 - \frac{x_i}{x}, & \text{if } x \in [x_i, 1); \\ 1, & \text{if } x = 1, \end{cases}$$

where x_i is determined by buyer i 's mean constraint.

¹⁷However, the existence or the exact shape of an asymmetric buyer-optimal information structure remains unknown to us at this time.

Proof. See Appendix A.11. □

Thus, both buyers have the virtual values $\{0, 1\}$ and the degree of asymmetry between p_1 and p_2 is reflected only by x_1 and x_2 . For $n \geq 3$, the seller-worst problem remains an isoperimetric problem and we can similarly reduce it to a finite-dimensional constrained optimization problem as we do in Theorem 1. The finite-dimensional problem, however, becomes intractable and its closed-form solution(s) remain unknown to us.

6.2 Continuous prior distributions

We now consider continuous prior distributions. Our analysis in Sections 3–5 remains applicable, if the information designer is allowed to choose any profile of signal distributions with a given mean p . However, if the information designer has full information about the prior, then a signal distribution is feasible if and only if it is a mean-preserving contraction of the prior; see Blackwell (1953). That is, the set of feasible signal distributions becomes

$$\mathcal{G}_H = \left\{ G : [0, 1] \rightarrow [0, 1] \left| \int_0^1 x dG(x) = p, \int_0^x G(t) dt \leq \int_0^x H(t) dt, \forall x \in [0, 1] \right. \right\}.$$

The main issue here is to handle the mean-preserving spread constraint on the signal distributions. Changing the variable is no longer useful so we need a different tool akin to the method developed in Dworczak and Martini (2019), although our objective function is still different from that in Dworczak and Martini (2019). Consider, for instance, the case of a second-price auction (with no reserve). The seller-worst problem can be expressed as

$$\begin{aligned} \max_G \int_0^1 (nG^{n-1}(x) - (n-1)G^n(x) - 1) dx \\ \text{s.t. } H \text{ is a mean-preserving spread of } G. \end{aligned} \tag{23}$$

We obtain the following result for the case with two buyers:

Proposition 3. *For $n = 2$, full revelation (i.e., $G = H$) solves the problem in (23).*

Proof. The objective is

$$\int_0^1 2G - G^2 - 1 dx = - \int_0^1 x d(2G - G^2) = \int_0^1 x dG^2(x) - 2p.$$

Moreover, $\int_0^1 x dG^2(x)$ is maximized if $G = H$, since G^2 is the CDF of the convex function, $\max\{x_1, x_2\}$, when x_1 and x_2 are independently distributed according to G . Hence, full revelation minimizes the sellers' revenue. □

Proposition 3 has an implication for both the seller-worst problem and the buyer-optimal problem with two buyers. Specifically, we can still argue that the seller-worst signal distribution must be regular and symmetric, and must admit only nonnegative virtual values. Since a second-price auction with

a reserve price is optimal for such signal distributions, identifying a seller-worst information structure amounts to solving the problem in (23). If the prior distribution is regular and admits nonnegative virtual values, it is also a feasible choice in this problem. Hence, the proof of Proposition 3 demonstrated that the full revelation is the seller-worst information structure. Moreover, since full revelation maximizes the total surplus, it is also the buyer-optimal information structure. The following corollary summarizes our findings.¹⁸

Corollary 2. *For $n = 2$, if the prior distribution is regular and admits nonnegative virtual values, then full revelation is both the unique symmetric buyer-optimal and the unique symmetric seller-worst information structure.*

Corollary 2 restores the equivalence between the buyer-optimal information structure and the seller-worst information structure. However, it requires the prior to be regular and induce only nonnegative virtual values. In particular, it rules out the binary prior which we analyze in Theorems 1 and 2.¹⁹ In Appendix A.13, we construct a regular continuous prior under which the buyer-optimal and seller-worst information structures are not equivalent. The example clarifies that the inequivalence of the seller-worst information and the buyer-optimal information arises not because the prior is discrete but rather it induces negative virtual values.

The uniqueness of the buyer-optimal information structure in Corollary 2 also contrasts with the multiplicity of buyer-optimal information structures available with a single buyer.²⁰ Since the information designer selects only signal distributions which are regular and admit nonnegative virtual values, the seller adopts a second-price auction. When $n = 2$, the objective function in (23) is strictly convex in G ; hence, full revelation is the unique buyer-optimal information structure. When $n = 1$, on the other hand, the objective becomes linear in G . Hence, any signal distribution that induces (i) the same posted price and (ii) the same probability of exceeding the posted price of a buyer-optimal signal distribution is also buyer-optimal.

¹⁸Based on our result that the seller-worst information is regular and induces nonnegative virtual value, Corollary 2 can alternatively be obtained from Part (i) in Theorem 5 of [Ganuzza and Penalva \(2010\)](#), which shows that in a second-price auction with two bidders, the seller's revenue is nonincreasing in the precision of the signals.

¹⁹If $n \geq 3$ or the prior is irregular, we can still solve the problem in (23); we report the solution in [Chen and Yang \(2021\)](#). However, the optimal signal distributions which we obtain in these situations are no longer regular. As we demonstrate in Example A.12, the optimal auction need not be a second-price auction with a reserve price.

²⁰For example, when $n = 1$ and the prior is uniformly distributed on $[1/2, 1]$, the buyer-optimal information structure identified in [Roesler and Szentes \(2017\)](#) is a truncated Pareto distribution $G(x) = 1 - \frac{1}{2x}$ for $x \in [1/2, 0.824)$ and $G(x) = 1$ for $x \in [0.824, 1]$. Against both the truncated Pareto distribution as well as the prior, the seller chooses the posted price $1/2$ and obtains revenue $1/4$.

6.3 Tightness of the seller-worst upper bound

As we mentioned in the introduction, our seller-worst revenue provides an upper bound for the revenue guarantee of any mechanism over all symmetric independent information structures and undominated equilibria. To illustrate the effectiveness of this upper bound, we consider a specific example with prior mean $p = 0.5$ and two buyers. It follows from Theorem 1 that the seller-worst revenue is $2x_s - x_s^2 = 0.3385$, where $x_s \approx 0.1867$ solves $x_s - x_s \log(x_s) = p$.

We now explore the revenue guaranteed by a second-price auction with a random reserve price distributed under the truthful equilibrium. This is consistent with the approach as taken by Che (2022), Park (2021), and Zhang (2021) in their study of the optimal revenue guarantee by a dominant-strategy mechanism with a mean constraint.²¹ Formally, denote by R the distribution function of a random reserve price. Let σ^T be the truth-telling equilibrium, i.e., $\sigma_i^T(x_i) = x_i$ for each x_i ; in addition, let $\Pi(R, G, \sigma^T)$ denote the seller's revenue under σ^T . The revenue guaranteed by R (under the truthful equilibrium) is defined as $\min_G \Pi(R, G, \sigma^T)$.

First, we examine the performance of a deterministic reserve price. In particular, Suzdaltsev (2020) proves that $r = 0$ solves

$$\max_r \min_G \Pi(r, G, \sigma^T).$$

It follows from Proposition 3 that the prior/full revelation solves $\min_G \Pi(0, G, \sigma^T)$. As a result, $\min_G \Pi(0, G, \sigma^T) = p^2 = 0.25$. Hence, any deterministic reserve guarantees the revenue 0.25, which is about 73.85 percent of the seller-worst revenue. In fact, by the same argument, for any p , any deterministic reserve guarantees at most $p^2/(2x_s - x_s^2)$ of the seller-worst revenue.

Second, we examine the performance of a specific random reserve price. Consider

$$R_b(r) = \begin{cases} 0, & \text{if } r \in [0, e^{-1/b}); \\ 1 + b \times \log(r), & \text{if } r \in [e^{-1/b}, 1]. \end{cases}$$

In particular, R_b becomes the random posted price due to Carrasco, Luz, Kos, Messner, Monteiro, and Moreira (2018) for $b = \frac{1}{-\log x_s} \approx 0.5958$.²² Our simulation result shows that when we choose $b = 0.5958$, we have $\min_G \Pi(R_b, G, \sigma^T) = 0.2939$, which is about 86.8 percent of the seller-worst revenue. In contrast, when $b = 0.339$, we have $\min_G \Pi(R_b, G, \sigma^T) = 0.3382$, which is about 99.9 percent of the seller-worst revenue.

²¹Park (2021) studies the optimal revenue guarantee in a public good provision setting, Zhang (2021) studies the problem in a bilateral trade setting, and Che (2022) studies the problem in an auction setting where mechanisms satisfy a condition called “competitiveness”. These papers all adopt a duality approach and allow for correlated signal distributions. In our example, the guaranteed revenue in Che (2022) is 0.317 which is about 93.6 percent of our seller-worst revenue. The optimal revenue guarantee in general dominant-strategy auctions remains unknown to us.

²²Carrasco, Luz, Kos, Messner, Monteiro, and Moreira (2018) proves that a random posted price attains the seller-worst revenue when there is a single buyer.

Third, [Bachrach, Chen, Talgam-Cohen, Yang, and Zhang \(2022\)](#) has identified, for $n = 2$ and any prior mean p , a random reserve price distribution $R^*(r)$ which guarantees the seller-worst revenue $2x_s - x_s^2$.²³ In particular,

$$R^*(r) = \begin{cases} \frac{r(1-x_s)}{r-x_s} \left(1 + \frac{1}{-\log(x_s)} \times \log(r)\right), & \text{if } r \in (0, x_s) \cup (x_s, 1]; \\ \frac{1-x_s}{-\log(x_s)}, & \text{if } r = x_s; \\ 0, & \text{if } r = 0. \end{cases}$$

7 Conclusion

In this paper, we characterize the symmetric buyer-optimal information structure as well as the symmetric seller-worst information structure with symmetric binary priors and a Myersonian optimal auction. We show that with a binary i.i.d. prior on 0 and 1, the two information structures are not equivalent, and yet both converge to “no disclosure” when the number of buyers goes to infinity. We also demonstrate that an asymmetric information structure is never seller-worst but can generate a strictly higher surplus for the buyers on an aggregate level.

The independent private-value setting enables us to express both the buyers’ surplus as well as the seller’s revenue in terms of the buyers’ (ironed) virtual values. This approach thereby neatly subsumes the IC and IR constraints and leads to an information design problem amenable to optimal control. For general correlated signal distributions, however, we know of no way to express the seller’s optimal revenue or the buyers’ surplus with the intractable (binding) IC and IR constraints.²⁴ One way to bypass this difficulty is to appeal to the strong duality approach used in [Du \(2018\)](#) and [Brooks and Du \(2021\)](#). That approach requires identifying a seller-worst/minmax information structure (among all correlated signal distributions) together with a maxmin mechanism that achieves the seller-worst revenue upper bound. We leave this important yet challenging question for future research.

Do the optimal information structures that we have provided resemble any real-world information structures? We do not have an answer. As information structures are inherently harder to observe than, say, contracts or selling mechanisms, we must also maintain the awareness that the predictions we have derived, like some of those in contract theory, might be entirely counterfactual. As we have demonstrated, however, the optimal information structures do provide useful theoretical benchmarks which shed light on other problems such as strategic information acquisitions or optimal revenue guarantees.

²³[Bachrach and Talgam-Cohen \(2022\)](#) incorporates material from these three other recent independent working papers: [Bachrach and Talgam-Cohen \(2022\)](#), [Chen and Yang \(2022\)](#), and [Zhang \(2022\)](#).

²⁴[Mathevet, Perego, and Taneva \(2020\)](#) study the information design problem allowing for correlated signals under a fixed game, whereas the game in our information design problem is chosen optimally by the seller in response to the information structure.

References

- BACHRACH, N., Y.-C. CHEN, I. TALGAM-COHEN, X. YANG, AND W. ZHANG (2022): “Distributionally Robust Auction Design,” *Working paper*.
- BACHRACH, N., AND I. TALGAM-COHEN (2022): “Distributional Robustness: From Pricing to Auctions,” *Working paper*.
- BERGEMANN, D., B. A. BROOKS, AND S. MORRIS (2016): “Informationally Robust Optimal Auction Design,” *Working paper*.
- BERGEMANN, D., AND M. PESENDORFER (2007): “Information Structures in Optimal Auctions,” *Journal of Economic Theory*, 137, 580–609.
- BLACKWELL, D. (1953): “Equivalent Comparisons of Experiments,” *The Annals of Mathematical Statistics*, 24, 265–272.
- BROOKS, B., AND S. DU (2021): “Optimal Auction Design with Common Values: An Informationally Robust Approach,” *Econometrica*, 89(3), 1313–1360.
- CARRASCO, V., V. F. LUZ, N. KOS, M. MESSNER, P. MONTEIRO, AND H. MOREIRA (2018): “Optimal Selling Mechanisms under Moment Conditions,” *Journal of Economic Theory*, 177, 245–279.
- CHE, E. (2022): “Distributionally Robust Optimal Auction Design under Mean Constraints,” *Working paper*.
- CHEN, X., Z. OWEN, C. PIXTON, AND D. SIMCHI-LEVI (2022): “A Statistical Learning Approach to Personalization in Revenue Management,” *Management Science*, 68(3), 1923–1937.
- CHEN, Y.-C., AND X. YANG (2021): “Information Design with Rank-Dependent Utilities,” *Working paper*.
- (2022): “Note on Revenue Guarantee for Second-Price Auctions with Random Reserve,” *Working paper*.
- DU, S. (2018): “Robust Mechanisms under Common Valuation,” *Econometrica*, 86, 1569–1588.
- DWORCZAK, P., AND G. MARTINI (2019): “The Simple Economics of Optimal Persuasion,” *Journal of Political Economy*, 127, 1993–2048.
- GANUZA, J.-J., AND J. S. PENALVA (2010): “Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions,” *Econometrica*, 78(3), 1007–1030.
- KRÄHMER, D. (2020): “Information Disclosure and Full Surplus Extraction in Mechanism Design,” *Journal of Economic Theory*, 187, 105020.

- KRISHNA, V. (2009): *Auction Theory*. Academic Press.
- MATHEVET, L., J. PEREGO, AND I. TANEVA (2020): “On Information Design in Games,” *Journal of Political Economy*, 128(4), 1370–1404.
- MONTEIRO, P. K. (2015): “A Note on the Continuity of the Optimal Auction,” *Economics Letters*, 137, 127–130.
- MONTEIRO, P. K., AND B. F. SVAITER (2010): “Optimal Auction with a General Distribution: Virtual Valuation without Densities,” *Journal of Mathematical Economics*, 46, 21–31.
- MYERSON, R. B. (1981): “Optimal Auction Design,” *Mathematics of Operations Research*, 6, 58–73.
- PARK, J. (2021): “Informational Regulation of Public Good Monopolist,” *Working paper*.
- ROESLER, A.-K., AND B. SZENTES (2017): “Buyer-Optimal Learning and Monopoly Pricing,” *American Economic Review*, 107, 2072–80.
- SHI, X. (2012): “Optimal Auctions with Information Acquisition,” *Games and Economic Behavior*, 74(2), 666–686.
- SUZDALTSEV, A. (2020): “Distributionally Robust Pricing in Independent Private Value Auctions,” *Working paper*.
- TERSTIEGE, S., AND C. WASSER (2020): “Buyer-Optimal Extensionproof Information,” *Journal of Economic Theory*, 188, 105070.
- VAN BRUNT, B. (2004): *The Calculus of Variations*. Springer.
- YANG, K. H. (2018): “Buyer-Optimal Information with Nonlinear Technology,” *Working paper*.
- ZHANG, W. (2021): “Robust Bilateral Trade Mechanism with Known Expectations,” *Working paper*.
- (2022): “Information-Robust Optimal Auctions,” *Working paper*.

Appendix

A Omitted proofs

A.1 Formal definition of ironed virtual valuations

For any CDF G with $\text{supp}(G) \subset [0, 1]$, let $a = \inf\{x \in [0, 1] | G(x) > 0\}$, and define

$$\Psi(x|G) = \begin{cases} 0, & \text{if } x \in [0, a); \\ a - x(1 - G(x)), & \text{if } x \in [a, 1]. \end{cases}$$

Let $\Theta = \{(\alpha, \beta) \in \mathbb{R}^2 | \alpha + \beta G(x) \leq \Psi(x|G), \forall x \in [0, 1]\}$ and let

$$\Phi(x|G) = \sup\{\alpha + \beta G(x) | (\alpha, \beta) \in \Theta\},$$

where $\Phi(x|G)$ is called the convexification of Ψ under the G -quantile space.

We say that $w(x)$ is a sub-gradient of $\Phi(x|G)$ at $x \in [0, 1]$ if

$$\Phi(z|G) - \Phi(x|G) \geq w(x)(G(z) - G(x)), \quad \forall z \in [0, 1].$$

For each $x \in [0, 1]$, let $\partial\Phi(x|G)$ denote the set of sub-gradients of $\Phi(\cdot|G)$ at x . Finally, let

$$\hat{\varphi}(x|G) = \inf \partial\Phi(x|G).$$

Then, $\hat{\varphi}(x|G)$ is defined as the *ironed virtual valuation* induced by G .

A.2 Proof of Lemma 1

Proof. When $\alpha = 1$, we first consider the following information design problem:

$$\max_{G \in \mathcal{G}_H} \left(\int_{[0,1]^n} \max\{x_1, \dots, x_n\} \prod_{i=1}^n (dG(x_i)) - \int_{[0,1]^n} \sum_{i=1}^n (\hat{\varphi}(x_i|G)) q_i(x_i, x_{-i}|G) \prod_{i=1}^n (dG(x_i)) \right). \quad (24)$$

Denote by $V_1(G)$ and $V_2(G)$ the objective of problem (1) and problem (24) under the signal distribution G respectively.

First, since $\max\{x_1, \dots, x_n\}$ is continuous in x , the first term in (24) is continuous in G . Moreover, by Theorem 2 of [Monteiro \(2015\)](#), the expected revenue is a lower semicontinuous function in G . Hence, $V_2(G)$ is an upper semicontinuous function in G . Also since $\mathcal{G}_H$ is a closed subset of the set of Borel probability measures on $[0, 1]$, $\mathcal{G}_H$ is compact. Thus, by the extreme value theorem, an optimal solution of the problem in (24) exists. Let G^* be the optimal solution to the problem (24).

Second, for any signal distribution G which induces negative virtual values with positive probability, we take the same modified distribution $\tilde{G}^{\theta_0} \in \mathcal{G}_H^+$ as follows,

$$\tilde{G}^{\theta_0}(x) = \begin{cases} \theta_0, & \text{if } x \in [0, x_{\theta_0}); \\ 1 - \frac{x_0(1-G(x_0^-))}{x}, & \text{if } x \in [x_{\theta_0}, x_0); \\ G(x), & \text{if } x \in [x_0, 1], \end{cases}$$

where θ_0 denotes the mass on $x = 0$. Since the virtual value on x_0 is $0 \in [\hat{\varphi}(x_0^-|G), \hat{\varphi}(x_0^+|G)]$, by Lemma 3, when $x \in [x_{\theta_0}, x_0]$, we have $G(x) \geq \tilde{G}^{\theta_0}(x)$. If $\theta_0 = 0$, then $\tilde{G}^0$ first-order stochastically dominates $G(x)$, and thereby $\int_0^1 x d\tilde{G}^0 \geq \int_0^1 x dG(x)$. If $\theta_0 = G(x_0^-)$, then $x_{\theta_0} = x_0$. Hence, $\tilde{G}^{G(x_0^-)}$ is first-order stochastically dominated by $G(x)$ and $\int_0^1 x d\tilde{G}^{G(x_0^-)} \leq \int_0^1 x dG(x)$. Since $\int_0^1 x d\tilde{G}^{\theta_0}(x)$ is continuous and strictly decreasing in θ_0 , it follows from the intermediate-value theorem that there exists a unique $\theta_0 \in [0, G(x_0^-)]$ such that $\int_0^1 x d\tilde{G}^{\theta_0} = p$. Moreover, by Lemma 3, since G assigns positive probabilities on negative virtual values for signals in $(0, 1]$, we have $\theta_0 \in (0, G(x_0^-))$. Hence, $\tilde{G}^{\theta_0}$ is a feasible signal distribution with nonnegative virtual values except at 0; moreover, $\tilde{G}^{\theta_0}$ is a strict mean-preserving spread of G . First, since the first term of the objective is convex in signal profile, the expectation of the first term under $\tilde{G}^{\theta_0}$ is strictly greater than that under G . Meanwhile, since $\hat{\varphi}(x|\tilde{G}^{\theta_0}) = 0$ and the seller only allocates the good to a buyer with a non-negative virtual value, the expected virtual value is the same under G and $\tilde{G}^{\theta_0}$. Therefore, $V_2(\tilde{G}^{\theta_0}) > V_2(G)$. Hence, $G^* \in \mathcal{G}_H^+$.

Third, we claim that G^* also solves the problem (1). To see this, observe that for any G and any signal realization $(x_1, \dots, x_n)$, $\sum_{i=1}^n x_i q_i(x_i, x_{-i}|G) \leq \max\{x_1, \dots, x_n\}$. If $G \in \mathcal{G}_H^+$, then for any signal realization $(x_1, \dots, x_n)$, $\sum_{i=1}^n x_i q_i(x_i, x_{-i}|G) = \max\{x_1, \dots, x_n\}$. Hence,

$$V_1(G) \leq V_2(G) \text{ for any } G \in \mathcal{G}_H, \text{ and } V_1(G) = V_2(G) \text{ for any } G \in \mathcal{G}_H^+. \quad (25)$$

Finally,

$$\max_{G \in \mathcal{G}_H} V_1(G) \leq \max_{G \in \mathcal{G}_H} V_2(G) = V_2(G^*) = V_1(G^*),$$

where the first inequality and the third equality follow from (25) and the second equality follows from the definition of G^* . Hence, G^* solves the problem in (1). Hence an optimal solution exists in the problem (1). $\square$

A.3 Proof of Lemma 4

Proof. The case with $\alpha = 1$: Let G be a signal distribution which assigns positive probability on negative virtual values for signals in $(0, 1]$. Let $x_0 = \inf\{x|\hat{\varphi}(x|G) \geq 0\}$. Define $\tilde{G}^{\theta_0}(x)$ such that

$$\tilde{G}^{\theta_0}(x) = \begin{cases} \theta_0, & \text{if } x \in [0, x_{\theta_0}); \\ 1 - \frac{x_0(1-G(x_0^-))}{x}, & \text{if } x \in [x_{\theta_0}, x_0); \\ G(x), & \text{if } x \in [x_0, 1], \end{cases}$$

where θ_0 denotes the mass on $x = 0$. Since the virtual value on x_0 is $0 \in [\hat{\varphi}(x_0^-|G), \hat{\varphi}(x_0^+|G)]$, by Lemma 3, when $x \in [x_{\theta_0}, x_0]$, we have $G(x) \geq \tilde{G}^{\theta_0}(x)$. If $\theta_0 = 0$, then $\tilde{G}^0$ first-order stochastically dominates $G(x)$, and thereby $\int_0^1 x d\tilde{G}^0 \geq \int_0^1 x dG(x)$. If $\theta_0 = G(x_0^-)$, then $x_{\theta_0} = x_0$. Hence, $\tilde{G}^{G(x_0^-)}$ is first-order stochastically dominated by $G(x)$ and $\int_0^1 x d\tilde{G}^{G(x_0^-)} \leq \int_0^1 x dG(x)$. Since $\int_0^1 x d\tilde{G}^{\theta_0}(x)$ is continuous and strictly decreasing in θ_0 , it follows from the intermediate-value theorem that there exists

a unique $\theta_0 \in [0, G(x_0^-)]$ such that $\int_0^1 x d\tilde{G}^{\theta_0} = p$. Moreover, by Lemma 3, since G assigns positive probabilities on negative virtual values for signals in $(0, 1]$, we have $\theta_0 \in (0, G(x_0^-))$.

Hence, $\tilde{G}^{\theta_0}$ is a feasible signal distribution with nonnegative virtual values except at 0; moreover, $\tilde{G}^{\theta_0}$ is a strict mean-preserving spread of G . Therefore, by Lemma 2, the total surplus under $\tilde{G}^{\theta_0}$ is strictly greater than the total surplus under G . In addition, since $\hat{\varphi}(x|\tilde{G}^{\theta_0}) = 0$, the seller only allocates the good to a buyer with a non-negative virtual value, the expected virtual value is the same under G and $\tilde{G}^{\theta_0}$. Hence, the objective value will be strictly higher as the expected total surplus becomes strictly higher and the seller's revenue remains the same. Hence, any optimal signal distribution G must induce nonnegative virtual values with probability one on $(0, 1]$.

The case with $\alpha = 0$: Suppose that $\tilde{G}^{\theta_0}$ puts some positive mass on $x = 0$. For $\alpha = 0$, we can further modify the distribution $\tilde{G}^{\theta_0}$ to reduce the seller's revenue as follows. Define another signal distribution,

$$\hat{G}^{x_1}(x) = \begin{cases} 1 - \frac{1 - \tilde{G}^{\theta_0}(x_1)}{x}, & \text{if } x \in [1 - \tilde{G}^{\theta_0}(x_1), x_1]; \\ \tilde{G}^{\theta_0}(x), & \text{if } x \in (x_1, 1]. \end{cases}$$

For $x_1 = x_0$, we have $\hat{G}^{x_1}(x) = \tilde{G}^{\theta_0}(x)$ which first-order stochastically dominates $\tilde{G}^{\theta_0}$, then $\int_0^1 x d\hat{G}^{x_0} \geq \int_0^1 x d\tilde{G}^{\theta_0}$, and for $x_1 = 1$, it is a degenerate distribution with all mass on $x = 0$. Also since $\int_0^1 x d\hat{G}^{x_1}$ is continuous and strictly decreasing in x_1 , the intermediate-value theorem implies that there exists a unique $x_1 \in (x_0, 1)$ such that $\int_0^1 x d\hat{G}^{x_1} = p$. Thus, $\hat{G}^{x_1}$ is a feasible information structure.

Since $\hat{\varphi}(x_1|\hat{G}^{x_1}) = 0 \leq \hat{\varphi}(x_1|\tilde{G}^{\theta_0})$, for each realization of signals, we have $\max\{0, \hat{\varphi}(x|\hat{G}^{x_1})\} \leq \max\{0, \hat{\varphi}(x|\tilde{G}^{\theta_0})\}$. Moreover, the inequality is strict with positive probability. Hence, the seller's revenue is strictly lower under $\hat{G}^{x_1}$ than under $\tilde{G}^{\theta_0}$. $\square$

A.4 Proof of Lemma 5

Proof. For any irregular distribution G with nonnegative virtual values except at 0, there exists some $x' > 0$ such that $\Psi(x'|G) < \Phi(x'|G)$. Since $G(x)$ is right continuous with respect to x' , so is $\Psi(x'|G)$. Therefore, there exists an interval $[x', x'']$, such that for any $x \in [x', x'']$, $\Psi(x|G) < \Phi(x|G)$. Let $[x_1, x_2] \supseteq [x', x'']$ be an ironed interval such that $\hat{\varphi}(x|G) = k$ is constant for $x \in [x_1, x_2]$. And for $x \in (x_1, x_2)$, $\Psi(x|G) \leq \Phi(x|G)$. For $x = x_1$ and $x = x_2$, $\Psi(x|G) = \Phi(x|G)$. Moreover, since $\Psi(x_1|G) = \Phi(x_1|G)$, $G(x)$ is continuous at x_1 . Then let $\hat{G}$ be

$$\hat{G} = \begin{cases} G(x), & \text{if } x \notin [x_1, x_2]; \\ 1 - \frac{(1 - G(x_1))(x_1 - k)}{x - k}, & \text{if } x \in [x_1, x_2]. \end{cases}$$

The modified distribution $\hat{G}$ has two key features: firstly, it generates the same virtual value as G for any realized signal x ; secondly, by lemma 3, since we have $\hat{G}(x) \leq G(x)$ and $\hat{G}(x) < G(x)$ on $x \in (x_1, x_2)$ (otherwise, $G(x)$ is regular), $\hat{G}$ will strictly first-order stochastically dominate $G(x)$, which implies that $\hat{G}$ will generate a strictly higher mean than p .

Hence, we can modify $\hat{G}$ again to satisfy the mean constraint. Let $\hat{G}^{\theta_0}$ be

$$\hat{G}^{\theta_0} = \begin{cases} \theta_0, & \text{if } x \in [0, x_{\theta_0}); \\ \hat{G}(x), & \text{if } x \in [x_{\theta_0}, 1], \end{cases}$$

where $\theta_0 = \hat{G}(x_{\theta_0})$ denotes the mass on signal $x = 0$.

It is clear that

$$\int_0^1 x d\hat{G}^{\theta_0}(x) = \int_0^1 x d\hat{G}(x) > p > 0 = \int_0^1 x d\hat{G}^1(x).$$

Since $\int_0^1 x d\hat{G}^{\theta_0}(x)$ is continuous and strictly decreasing in θ_0 , the intermediate-value theorem implies that there exists a unique $\theta_0 \in (0, 1)$ such that $\int_0^1 x d\hat{G}^{\theta_0} = p$. Therefore, $\hat{G}^{\theta_0}$ is a feasible signal distribution.

By construction,

$$\hat{\varphi}(x|\hat{G}^{\theta_0}) = \begin{cases} -\frac{1-\theta_0}{\theta_0} < 0, & \text{if } x \in [0, x_{\theta_0}); \\ \hat{\varphi}(x|\hat{G}) = \hat{\varphi}(x|G), & \text{if } x \in [x_{\theta_0}, 1]. \end{cases}$$

Hence, $\max\{\hat{\varphi}(x|\hat{G}^{\theta_0}), 0\} \leq \max\{\hat{\varphi}(x|G), 0\}$.

Hence, given any signal distribution G , we can modify a buyer's signal distribution to $\hat{G}^{\theta_0}$. As the arguments in Lemma 4, this modification has two effects: first, the seller's revenue (the second term of the objective function) as a max function of nonnegative virtual values becomes weakly less; second, since the constructed distribution $\hat{G}^{\theta_0}$ is a strict mean-preserving spread of the original distribution G , by Lemma 4 and Lemma 2, the total surplus is strictly higher. Therefore, with this modification, the buyers' total surplus is strictly higher. Hence, the buyer-optimal distribution must be regular except at 0.

Moreover, this modification still puts a positive mass on signal 0. By Lemma 4, we can further modify this distribution into another distribution with nonnegative virtual values and generate strictly less revenue for the seller. Thus, the seller-worst distribution must also be regular. $\square$

A.5 Change of variable

Denote $G_m(0)$ by $F_m(0^-)$, we have

$$\begin{aligned} \int_0^1 x dG_m^n(x) &= \int_0^1 \left(k + \frac{\int_k^1 (1 - F_m(s)) ds}{1 - F_m(k)} \right) dF_m^n(k) \\ &= \int_0^1 k dF_m^n(k) + n \int_0^1 \left(\frac{\int_k^1 (1 - F_m(s)) ds}{1 - F_m(k)} \right) F_m^{n-1}(k) dF_m(k) \\ &= \int_0^1 k dF_m^n(k) + n \int_0^1 \int_k^1 (1 - F_m(s)) \frac{F_m^{n-1}(k)}{1 - F_m(k)} ds dF_m(k) \\ &= \int_0^1 k dF_m^n(k) + n \int_0^1 (1 - F_m(s)) \int_0^s \frac{F_m^{n-1}(k)}{1 - F_m(k)} dF_m(k) ds \\ &= \int_0^1 k dF_m^n(k) + n \int_0^1 (1 - F_m(s)) \int_{F_m(0^-)}^{F_m(s)} \frac{\theta^{n-1}}{1 - \theta} d\theta ds \\ &= \int_0^1 n(1 - F_m(k)) \left(\sum_{i=1}^{n-1} \frac{-F_m^i(k)}{i} - \log(1 - F_m(k)) + \left(\sum_{i=1}^{n-1} \frac{F_m^i(0^-)}{i} + \log(1 - F_m(0^-)) \right) \right) - F_m^n(k) dk + 1. \end{aligned}$$

Then by bounded convergence theorem, since G_m and F_m uniformly converge to G and F respectively, Equation (16) holds.

A.6 Proof of Lemma 9

Proof. First, define

$$\zeta_\alpha(\theta) = n\theta^{n-2}(n-1-(n-1+\alpha)\theta) = (1-\theta)I'_\alpha(\theta) - \lambda.$$

We discuss the cases with $\alpha = 0$ and $\alpha = 1$, respectively.

1. For $\alpha = 0$, we have $\zeta_0(\theta) = n(n-1)(1-\theta)\theta^{n-2}$, then

$$\frac{\partial \zeta_0(\theta)}{\partial \theta} = n(n-1)\theta^{n-3}(n-2-(n-1)\theta).$$

Therefore, $\zeta_0(\theta)$ is increasing in θ when $\theta \in (0, (n-2)/(n-1))$, and then decreasing in θ when $\theta \in ((n-2)/(n-1), 1)$. Moreover, we have $\zeta_0(0) = \zeta_0(1) = 0$. We then have the following three cases:

- (a) *Case 1.* $\lambda \geq 0$: In this case, $I'_0(\theta)$ is always positive for both $\alpha = 0$. Hence, $I_0(\theta)$ is increasing. Thus, $I_0(\theta)$ will cross the θ -axis from below at most once. The objective function takes a local minimal and hence we ignore this case.
- (b) *Case 2.* $\lambda \in [\lambda_0^*, 0)$ where $\lambda_0^* = -\zeta_0((n-2)/(n-1)) < 0$: $I'_0(\theta)$ is first negative, then positive, and eventually becomes negative. In addition, $I_0(\theta) = 0$ when $\theta = \theta_0$. Hence $I_0(\theta)$ will first decrease from 0, then increase, and finally decrease. In this case, $I_0(\theta)$ will cross the θ -axis at most twice. Only for the second time, $I_0(\theta)$ will cross the θ -axis from above.
- (c) *Case 3.* $\lambda < \lambda_0^*$: Then $I_0(\theta)$ is always decreasing. Thus, $I_0(\theta)$ crosses the θ -axis from above at most once.

2. For $\alpha = 1$, we have $\zeta_1(\theta) = n\theta^{n-2}(n-1-n\theta)$, then

$$\frac{\partial \zeta_1(\theta)}{\partial \theta} = n\theta^{n-3}((n-2)(n-1-n\theta) - n\theta) = n(n-1)\theta^{n-3}((n-2) - n\theta).$$

Therefore, $\zeta_1(\theta)$ is increasing in θ when $\theta \in (0, (n-2)/n)$, and then decreasing in θ when $\theta \in ((n-2)/n, 1)$. Moreover, we have $\zeta_1(0) = 0$ and $\zeta_1(1) = -n$. We then have the following four cases:

- (a) *Case 1.* $\lambda \geq n$: In this case, $I'_1(\theta)$ is always positive. Hence, $I_1(\theta)$ is increasing. Thus, $I_1(\theta)$ will cross the θ -axis from below at most once. The objective function takes a local minimal and hence we ignore the case.
- (b) *Case 2.* $\lambda \in [0, n)$: $I'_1(\theta)$ is first positive and then negative, therefore, $I_1(\theta)$ will first increase and then decrease. In this case, $I_1(\theta)$ will cross the θ -axis at most twice. However, only for the second time, $I_1(\theta)$ will cross the θ -axis from above. Hence there is only one θ such that $I_1(\theta) = 0$ and also satisfying the second-order condition.

- (c) *Case 3.* $\lambda \in [\lambda_1^*, 0)$ where $\lambda_1^* = -\zeta_1((n-2)/n) < 0$: $I_1'(\theta)$ is first negative, then positive, and eventually becomes negative. In addition, $I_1(\theta) = 0$ when $\theta = \theta_0$. Hence $I_1(\theta)$ will first decrease from 0, then increase, and finally decrease. In this case, $I_1(\theta)$ will cross the θ -axis at most twice. Only for the second time, $I_1(\theta)$ will cross the θ -axis from above.
- (d) *Case 4.* $\lambda < \lambda_1^*$: Then $I_1(\theta)$ will be always decreasing. Thus $I_1(\theta)$ will also cross the θ -axis from above at most once.

□

A.7 Finite-dimensional seller-worst optimization

To obtain the solution, we solve Problem (19) with $\alpha = 0$:

$$\begin{aligned} \max_{k \geq 0, \theta} & -(\theta^n \times k + (1 - \theta^n) \times 1) \\ \text{s.t. } & k + (1 - k)(1 - \theta)(1 - \log(1 - \theta)) = p. \end{aligned}$$

Given the Lagrangian multiplier λ_s , the Lagrangian is

$$\mathcal{L}_s(\theta, k, \lambda_s) = (1 - k)\theta^n - \lambda_s(k + (1 - k)(1 - \theta)(1 - \log(1 - \theta))) + \lambda_s p - 1.$$

The Euler-Lagrange condition with respect to θ implies:

$$\frac{\partial \mathcal{L}_s}{\partial \theta} = (1 - k)(n\theta^{n-1} - \lambda_s \log(1 - \theta)) = 0.$$

Hence,

$$\lambda_s = \frac{n\theta^{n-1}}{\log(1 - \theta)}. \quad (26)$$

Recall

$$J_s(\theta) = (\theta \log(1 - \theta) + n(\theta + (1 - \theta) \log(1 - \theta))).$$

Then, taking the derivative of $\mathcal{L}_s$ with respect to k and using (26), we obtain

$$\begin{aligned} \frac{\partial \mathcal{L}_s}{\partial k} &= -\theta^n - \lambda_s(1 - (1 - \theta)(1 - \log(1 - \theta))) \\ &= \frac{\theta^{n-1}}{-\log(1 - \theta)} (\theta \log(1 - \theta) + n(\theta + (1 - \theta) \log(1 - \theta))) = \frac{\theta^{n-1}}{-\log(1 - \theta)} J_s(\theta). \end{aligned}$$

Hence, the sign of $\frac{\partial \mathcal{L}_s}{\partial k}$ is determined by the sign of $J_s(\theta)$. Moreover, we have

$$\begin{aligned} J_s'(\theta) &= \log(1 - \theta) - \frac{\theta}{1 - \theta} + n - n - n \log(1 - \theta) = -(n - 1) \log(1 - \theta) - \frac{\theta}{1 - \theta}, \\ J_s''(\theta) &= \frac{(n - 2) - (n - 1)\theta}{(1 - \theta)^2}. \end{aligned}$$

- When $n = 2$, for $\theta > 0$, we have $J_s''(\theta) < 0$; therefore, $J_s'(\theta) < J_s'(0) = 0$. Hence $J_s(\theta)$ is always less than $J_s(0) = 0$. As a result, the optimal $k_s = 0$ and we set $p_s = 1$.

- When $n \geq 3$, $J_s''(\theta) = 0$ has a unique solution $\theta_1 = (n-2)/(n-1)$. Moreover, $J_s'(\theta)$ is increasing in θ if $\theta \in (0, \theta_1)$ and decreasing in θ if $\theta \in (\theta_1, 1)$. Since $1/(1-\theta)$ decreases faster than $\log(1-\theta)$, $\lim_{\theta \uparrow 1} J_s'(\theta) \rightarrow -\infty$. Furthermore, since $J_s'(0) = 0$, there exists $\theta_2 \in (\theta_1, 1)$ such that for $\theta \in (0, \theta_2)$, $J_s'(\theta)$ is greater than 0 and for $\theta \in (\theta_2, 1)$, $J_s'(\theta)$ is less than 0. Therefore, $J_s(\theta)$ is increasing for $\theta \in (0, \theta_2)$ and decreasing for $\theta \in (\theta_2, 1)$. In addition, $J_s(0) = 0$ and $\lim_{\theta \uparrow 1} J_s(\theta) \rightarrow -\infty$. Therefore, there exists a unique $\theta_s \in (\theta_2, 1)$ such that $J_s(\theta_s) = 0$; moreover, $J_s(\theta) > 0$ if $\theta < \theta_s$ and $J_s(\theta) < 0$ if $\theta > \theta_s$. Recall that the threshold

$$p_s = (1 - \theta_s)(1 - \log(1 - \theta_s)),$$

where θ_s satisfies $J_s(\theta_s) = 0$. If $k_s > 0$ with the mass θ_s , the mean constraint requires $p > p_s$. Moreover, the corner solution occurs when $0 < p \leq p_s$. Notice that for an interior solution, θ_s only depends on n and so does p_s .

In summary, as stated in Theorem 1, the seller-worst information is the truncated Pareto distribution G_s in (3) parametrized by (k_s, x_s) .

- When $n = 2$, for any $p \in (0, 1)$, we have a corner solution $k_s = 0$. As a result, x_s is pinned down by the mean constraint $x_s(1 - \log(x_s)) = p$.
- When $n \geq 3$ and $p \in (0, p_s]$, we have a corner solution $k_s = 0$. As a result, x_s is pinned down by the mean constraint $x_s(1 - \log(x_s)) = p$.
- When $n \geq 3$ and $p \in (p_s, 1)$, k_s is an interior solution and $\theta_s > 0$ solves $J_s(\theta_s) = 0$; and k_s is pinned down by the mean constraint

$$(1 - k_s)(1 - \theta_s)(1 - \log(1 - \theta_s)) + k_s = p.$$

Finally, plugging θ_s and k_s into $\theta_s \equiv 1 - (x_s - k_s) / (1 - k_s)$, we obtain x_s .

A.8 Finite-dimensional buyer-optimal optimization

To obtain the solution, we solve Problem (19) with $\alpha = 1$:

$$\begin{aligned} \max_{\{\theta_0 \geq 0, k \geq 0, \theta\}} & n(1-k)(1-\theta) \left(\sum_{i=1}^{n-1} \frac{-\theta^i}{i} - \log(1-\theta) + \left(\sum_{i=1}^{n-1} \frac{\theta_0^i}{i} + \log(1-\theta_0) \right) \right) \\ \text{s.t.} & (1-k)((1-\theta)(1-\log(1-\theta) + \log(1-\theta_0)) + k(1-\theta_0)) = p. \end{aligned} \quad (27)$$

Given the Lagrangian multiplier λ_b , the Lagrangian is

$$\begin{aligned} \mathcal{L}_b(\theta_0, k, \theta, \lambda_b) = & n(1-k)(1-\theta) \left(\sum_{i=1}^{n-1} \frac{-\theta^i}{i} - \log(1-\theta) + \left(\sum_{i=1}^{n-1} \frac{\theta_0^i}{i} + \log(1-\theta_0) \right) \right) \\ & - \lambda_b ((1-k)(1-\theta)(1-\log(1-\theta) + \log(1-\theta_0)) + k(1-\theta_0)) + \lambda_b p. \end{aligned}$$

The Euler-Lagrange condition with respect to θ implies:

$$\begin{aligned}
\frac{\partial \mathcal{L}_b}{\partial \theta} &= (1-k) \left(-n \left(\sum_{i=1}^{n-1} \frac{-\theta^i}{i} - \log(1-\theta) + \left(\sum_{i=1}^{n-1} \frac{\theta_0^i}{i} + \log(1-\theta_0) \right) \right) + n(1-\theta) \left(\sum_{i=1}^{n-1} -\theta^{i-1} + \frac{1}{1-\theta} \right) \right) \\
&\quad - \lambda_b(1-k) (- (1 - \log(1-\theta) + \log(1-\theta_0)) + 1) \\
&= (1-k) \left(-n \left(\sum_{i=1}^{n-1} \frac{-\theta^i}{i} - \log(1-\theta) + \left(\sum_{i=1}^{n-1} \frac{\theta_0^i}{i} + \log(1-\theta_0) \right) \right) + n(1-\theta) \left(-\frac{1-\theta^{n-1}}{1-\theta} + \frac{1}{1-\theta} \right) \right) \\
&\quad - \lambda_b(1-k) (- (1 - \log(1-\theta) + \log(1-\theta_0)) + 1) \\
&= (1-k) \left(n \left(\theta^{n-1} + \left(\sum_{i=1}^{n-1} \frac{\theta^i}{i} + \log(1-\theta) \right) - \left(\sum_{i=1}^{n-1} \frac{\theta_0^i}{i} + \log(1-\theta_0) \right) \right) - \lambda(\log(1-\theta) - \log(1-\theta_0)) \right) = 0.
\end{aligned}$$

Thus, we have

$$\lambda_b = \frac{n \left(\theta^{n-1} + \left(\sum_{i=1}^{n-1} \frac{\theta^i}{i} + \log(1-\theta) \right) - \left(\sum_{i=1}^{n-1} \frac{\theta_0^i}{i} + \log(1-\theta_0) \right) \right)}{\log(1-\theta) - \log(1-\theta_0)}. \quad (28)$$

Also, taking the derivative of $\mathcal{L}_b$ with respect to θ_0 , we have

$$\begin{aligned}
\frac{\partial \mathcal{L}_b}{\partial \theta_0} &= (1-k)(1-\theta)n \left(\sum_{i=1}^{n-1} \theta_0^{i-1} - \frac{1}{1-\theta_0} \right) - \lambda_b \left(-\frac{(1-k)(1-\theta)}{1-\theta_0} - k \right) \\
&= \frac{-(1-k)(1-\theta)n\theta_0^{n-1}}{1-\theta_0} - \lambda_b \frac{-(1-k)(1-\theta) - k(1-\theta_0)}{(1-\theta_0)} \\
&= -\frac{(1-k)(1-\theta)n\theta_0^{n-1} - \lambda_b(1-\theta + k(\theta - \theta_0))}{(1-\theta_0)}. \quad (29)
\end{aligned}$$

Taking the derivative of $\mathcal{L}_b$ with respect to k , we have

$$\begin{aligned}
\frac{\partial \mathcal{L}_b}{\partial k} &= (1-\theta) \left(n \left(\sum_{i=1}^{n-1} \frac{\theta^i}{i} + \log(1-\theta) \right) - n \left(\sum_{i=1}^{n-1} \frac{\theta_0^i}{i} + \log(1-\theta_0) \right) \right) \\
&\quad + \lambda_b(1-\theta)(1 - \log(1-\theta) + \log(1-\theta_0)) - \lambda_b(1-\theta_0).
\end{aligned}$$

Since k and θ_0 may have corner solutions, we will study the corner solutions and interior solutions by cases. And we will show that k and θ_0 cannot be both interior.

- If $k > 0$, by Equation (28), we have

$$\begin{aligned}
&\lambda_b(1-\theta)(\log(1-\theta) + \log(1-\theta_0)) \\
&= n(1-\theta)\theta^{n-1} + n(1-\theta) \left(\left(\sum_{i=1}^{n-1} \frac{\theta^i}{i} + \log(1-\theta) \right) - \left(\sum_{i=1}^{n-1} \frac{\theta_0^i}{i} + \log(1-\theta_0) \right) \right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{\partial \mathcal{L}_b}{\partial k} &= \lambda_b(1-\theta)(\log(1-\theta) + \log(1-\theta_0)) - n(1-\theta)\theta^{n-1} + \lambda_b(1-\theta)(1 - \log(1-\theta) + \log(1-\theta_0)) - \lambda_b(1-\theta_0) \\
&= \lambda_b(1-\theta) - n(1-\theta)\theta^{n-1} - \lambda_b(1-\theta_0) = 0.
\end{aligned}$$

Hence, we have

$$\lambda_b = \frac{-n(1-\theta)\theta^{n-1}}{\theta - \theta_0}.$$

Since $\theta_b \geq \theta_0$, we have

$$\begin{aligned}\frac{\partial \mathcal{L}_b}{\partial \theta_0} &= \frac{(1-k)(1-\theta)n\theta_0^{n-1} + \lambda_b((1-\theta) + k(\theta - \theta_0))}{-(1-\theta_0)} \\ &= \frac{n(1-\theta)}{-(1-\theta_0)} \left((1-k)\theta_0^{n-1} + \frac{\theta^{n-1}(1-\theta)}{\theta - \theta_0} + k\theta^{n-1} \right) \\ &= \frac{n(1-\theta)}{-(1-\theta_0)} \left(\theta_0^{n-1} + \frac{\theta^{n-1}(1-\theta)}{\theta - \theta_0} + k(\theta^{n-1} - \theta_0^{n-1}) \right) < 0.\end{aligned}$$

Therefore, as long as $k > 0$, we have $\frac{\partial \mathcal{L}_b}{\partial \theta_0} < 0$, and thereby θ_0 will go down to 0 and be a *corner* solution.

We then pin down the support of p such that the optimal $k > 0$. Given $\theta_0 = 0$, let

$$\begin{aligned}J_b(\theta) &\equiv \log(1-\theta) + \theta^{n-2}(\theta + (1-\theta)\log(1-\theta)) + \sum_{i=1}^{n-1} \frac{\theta^i}{i} \\ &= \theta^{n-1} + \sum_{i=1}^{n-1} \frac{\theta^i}{i} + \log(1-\theta)(1 + (1-\theta)\theta^{n-2})\end{aligned}\tag{30}$$

We then have

$$\frac{\partial \mathcal{L}_b}{\partial k} = -\frac{n\theta \left(\theta^{n-1} + \left(\sum_{i=1}^{n-1} \frac{\theta^i}{i} + \log(1-\theta) \right) \right)}{\log(1-\theta)} - n(1-\theta)\theta^{n-1} = \frac{n\theta}{-\log(1-\theta)} J_b(\theta).$$

Hence, the sign of $\frac{\partial \mathcal{L}_b}{\partial k}$ is determined by the sign of $J_b(\theta)$.²⁵ Moreover, we have

$$\begin{aligned}J'_b(\theta) &= \frac{\theta^{n-3}(\theta + (1-\theta)\log(1-\theta))}{1-\theta} ((n-2) - (n-1)\theta), \\ \implies \text{Sign } J'_b(\theta) &= \text{Sign } ((n-2) - (n-1)\theta).\end{aligned}$$

As in the seller-worst case,

- When $n = 2$, for any $\theta > 0$, $J'_b(\theta) < 0$ and $J_b(\theta) < J_b(0) = 0$. As a result, the optimal $k_b = 0$ when $n = 2$. In this case, we set $p_b = 1$.
- When $n \geq 3$, for $\theta \in (0, (n-2)/(n-1))$, $J'_b(\theta) > 0$ and for $\theta \in ((n-2)/(n-1), 1]$, $J'_b(\theta) < 0$. In addition that $J_b(0) = 0$ and $\lim_{\theta \uparrow 1} J_b(\theta) \rightarrow -\infty$, there exists a unique $\theta_b \in (0, 1)$ such that $J_b(\theta_b) = 0$; moreover, $J_b(\theta) > 0$ if $\theta < \theta_b$ and $J_b(\theta) < 0$ if $\theta > \theta_b$.

Recall that the threshold

$$p_b = (1 - \theta_b)(1 - \log(1 - \theta_b)),$$

where θ_b satisfies $J_b(\theta_b) = 0$. If $k_b > 0$ with the mass θ_b , the mean constraint requires $p > p_b$.

Moreover, the corner solution $k_b = 0$ occurs when $0 < p \leq p_s$.

- If $k_b = 0$, we have $p \in (0, p_b]$. We then only need to solve θ_0 and θ . Moreover, if the optimal θ_0 is chosen, θ will be automatically pinned down by the mean constraint. We also have two cases for the optimal θ_0 : $\theta_0 > 0$ and $\theta_0 = 0$.

²⁵We present instead $\theta J_b(\theta)$ in the cost-benefit equation (10) for the ease of comparison with $J_s(\theta)$.

Suppose that $\theta_0 > 0$, then the Euler-Lagrange condition for θ_0 in Equation (29) implies

$$\begin{aligned}\frac{\partial \mathcal{L}_b}{\partial \theta_0} &= -\frac{(1-k_b)(1-\theta)n\theta_0^{n-1} - \lambda_b(1-\theta + k(\theta - \theta_0))}{(1-\theta_0)} \\ &= \underbrace{-\frac{(1-\theta)n\theta_0^{n-1}}{(1-\theta_0)}}_{\text{cost}(<0)} + \underbrace{\frac{\lambda_b(1-\theta)}{(1-\theta_0)}}_{\text{benefit}(>0)} = 0.\end{aligned}\quad (31)$$

Raising the mass θ_0 at signal 0 has two countervailing effects on the objective in (27). First, by increasing θ_0 , $\sum_{i=1}^{n-1} \frac{\theta_0^i}{i} + \log(1-\theta_0)$ in (27) decreases, which translates into a cost in proportion to the first term in (31). Second, to obey the mean constraint, the probability θ is reduced, thus $\sum_{i=1}^{n-1} \frac{\theta^i}{i} - \log(1-\theta)$ in (27) increases, which results in a benefit in proportion to the second term in (31).

Together with Equation (28), we have

$$\lambda_b = n\theta_0^{n-1} = \frac{n\left(\theta^{n-1} + \left(\sum_{i=1}^{n-1} \frac{\theta^i}{i} + \log(1-\theta)\right) - \left(\sum_{i=1}^{n-1} \frac{\theta_0^i}{i} + \log(1-\theta_0)\right)\right)}{\log(1-\theta) - \log(1-\theta_0)}.\quad (32)$$

We have if $\lambda_b > 0$, θ_0 is positive; if $\lambda_b \leq 0$, $\theta_0 = 0$. In order that there exists a positive solution (θ, θ_0) , we also have some restrictions on the mean p .

More precisely, there exists another threshold,²⁶

$$r_b = (1 - \theta_{r_b})(1 - \log(1 - \theta_{r_b})),$$

where θ_{r_b} satisfies Equation (32) with $\lambda_b = 0$ and $\theta_0 = 0$, that is,

$$\theta_{r_b}^{n-1} + \left(\sum_{i=1}^{n-1} \frac{\theta_{r_b}^i}{i} + \log(1 - \theta_{r_b})\right) = 0.$$

When $p \in (0, r_b)$, λ_b will be positive. In this case, θ_0 will be also positive. When $p \in [r_b, p_b]$, $\frac{\partial \mathcal{L}_b}{\partial \theta_0}$ is always less than 0, θ_0 will go down to 0 and both k_b and θ_0 will be corner solutions.

In summary, as stated in Theorem 2, the buyer optimal information ($\alpha = 1$) is the truncated Pareto distribution G_b in (8) parameterized by $(\theta_0, k_b, \theta_b)$.

- When $n = 2$, for any $p \in (0, 1)$, the optimal $k_b = 0$ is a corner solution.
 - When $p \in (0, r_b)$, θ_0 is an interior solution. (θ_0, θ_b) is jointly determined by the Euler-Lagrange condition and the mean constraint,

$$\begin{aligned}\theta_0 &= \frac{(2\theta_b + \log(1 - \theta_b)) - (\theta_0 + \log(1 - \theta_0))}{\log(1 - \theta_b) - \log(1 - \theta_0)}, \\ (1 - \theta_b)(1 - \log(1 - \theta_b) + \log(1 - \theta_0)) &= p.\end{aligned}$$

- When $p \in [r_b, 1)$, both $k_b = 0$ and $\theta_0 = 0$ are corner solutions. The only parameter x_b is then pinned down by the mean constraint $x_b(1 - \log(x_b)) = p$.

²⁶Note that $\theta_{r_b} > \theta_b$ and thereby $r_b < p_b$. This is because the factor $(1 - \theta)\theta^{n-2}$ in Expression (30) makes $J_b(\theta)$ arrives 0 faster; and the function $(1 - \theta)(1 - \log(1 - \theta))$ is decreasing in θ .

- When $n = 3$, there are two thresholds r_b and p_b .
 - When $p \in (p_b, 1)$, k_b is an interior solution and $\theta_0 = 0$ is a corner solution. $\theta_b > 0$ solves $J_b(\theta_b) = 0$; and k_b is pinned down by the mean constraint

$$(1 - k_b)(1 - \theta_b)(1 - \log(1 - \theta_b)) + k_b = p.$$

- When $p \in [r_b, p_b]$, both $k_b = 0$ and $\theta_0 = 0$ are corner solutions. The only parameter x_b is then pinned down by the mean constraint $x_b(1 - \log(x_b)) = p$.
- When $p \in (0, r_b)$, $k_b = 0$ is a corner solution and θ_0 is an interior solution. (θ_0, θ_b) is jointly determined by the Euler-Lagrange condition and the mean constraint,

$$\theta_0^{n-1} = \frac{\left(\theta_b^{n-1} + \left(\sum_{i=1}^{n-1} \frac{\theta_b^i}{i} + \log(1 - \theta_b)\right) - \left(\sum_{i=1}^{n-1} \frac{\theta_0^i}{i} + \log(1 - \theta_0)\right)\right)}{\log(1 - \theta_b) - \log(1 - \theta_0)},$$

$$(1 - \theta_b)(1 - \log(1 - \theta_b) + \log(1 - \theta_0)) = p.$$

Lastly, we show the following lemma that the benefit in (6) is greater than the benefit in (10). Hence, the buyer-optimal information designer is more reluctant to raise the low virtual value than a seller-worst information designer.

Claim 1. For any $n \geq 3$, the benefit in (6) is greater than the benefit in (10).

Proof. Define the difference of the benefit in (6) and the benefit in (10) by $\Delta J_n(\theta) \times \theta$. Hence, we have

$$\Delta J_n(\theta) = (n - \theta^{n-1}) \left(1 + \frac{(1 - \theta) \log(1 - \theta)}{\theta}\right) - \sum_{i=1}^{n-1} \frac{\theta^i}{i}.$$

We then define

$$\Delta J_{n2}(\theta) = (n - \theta^{2-1}) \left(1 + \frac{(1 - \theta) \log(1 - \theta)}{\theta}\right) - \sum_{i=1}^{n-1} \frac{\theta^i}{i}.$$

We have $\Delta J_n(\theta) \geq \Delta J_{n2}(\theta)$ for any $n \geq 3$. Then, taking the derivative of ΔJ_{n2} with respect to θ , we obtain

$$\frac{\partial \Delta J_{n2}}{\partial \theta} = \frac{n}{\theta^2} (-\log(1 - \theta) - \theta) + \log(1 - \theta) - \frac{1 - \theta^{n-1}}{1 - \theta}.$$

Moreover,

$$\begin{aligned} \frac{\partial^2 \Delta J_{n2}}{\partial \theta \partial n} &= \frac{-\log(1 - \theta) - \theta}{\theta^2} + \frac{\theta^{n-1} \log(\theta)}{1 - \theta} \\ &\geq \frac{\partial^2 \Delta J_{n2}}{\partial \theta \partial n} \Big|_{n=3} (\theta) > 0, \quad \forall \theta \in (0, 1). \end{aligned}$$

That is, $\frac{\partial \Delta J_{n2}}{\partial \theta}$ is increasing in n . Hence, to show that $\frac{\partial \Delta J_{n2}}{\partial \theta} > 0$ for any $n \geq 3$, it suffices to show $\frac{\partial \Delta J_{n2}}{\partial \theta} > 0$ when $n = 3$. First, we have

$$\frac{\partial^2 \Delta J_{n2}}{\partial \theta^2} \Big|_{n=3} = \frac{6 \log(1 - \theta) - \frac{(\theta - 2)\theta(\theta^2 - 3)}{\theta - 1}}{\theta^3}.$$

Then, we have

$$\frac{\partial}{\partial \theta} \left(6 \log(1 - \theta) - \frac{(\theta - 2)\theta(\theta^2 - 3)}{\theta - 1} \right) = \frac{\theta^2(-3\theta^2 + 8\theta - 3)}{(1 - \theta)^2}.$$

For $\theta \in (0, \frac{1}{3}(4 - \sqrt{7}))$, we have $-3\theta^2 + 8\theta - 3 < 0$ and for $\theta \in (\frac{1}{3}(4 - \sqrt{7}), 1)$, we have $-3\theta^2 + 8\theta - 3 > 0$.

Therefore, $\frac{\partial^2 \Delta J_{n2}}{\partial \theta^2} \Big|_{n=3}$ is first decreasing and then increasing. Since $\lim_{\theta \downarrow 0} \frac{\partial^2 \Delta J_{n2}}{\partial \theta^2} \Big|_{n=3}(\theta) = -1$ and $\lim_{\theta \uparrow 1} \frac{\partial^2 \Delta J_{n2}}{\partial \theta^2} \Big|_{n=3}(\theta) \rightarrow \infty$, $\frac{\partial \Delta J_{n2}}{\partial \theta} \Big|_{n=3}$ is also decreasing first and then increasing. Therefore, $\frac{\partial \Delta J_{n2}}{\partial \theta} \Big|_{n=3}$ has a unique minimum over $\theta \in (0, 1)$ which is strictly positive. Hence, when $n = 3$, $\frac{\partial \Delta J_{n2}}{\partial \theta} > 0$ for any θ . Thus, $\Delta J_n(\theta) > \Delta J_{n2}(\theta) \geq \Delta J_{n2}(0) = 0$. Therefore, the benefit in (6) is greater than the benefit in (10). $\square$

A.9 Proof of Lemma 11

Proof of nonnegativity. For each buyer i , let G_i be a signal distribution which assigns positive probability on negative virtual values for signals in $[0, 1]$ and let $x_{i0} = \inf\{x | \varphi(x|G_i) \geq 0\}$.

First, we construct a distribution $\tilde{G}_i^{\theta_{i0}}$ similar to that in Appendix A.3:

$$\tilde{G}_i^{\theta_{i0}}(x) = \begin{cases} \theta_{i0}, & \text{if } x \in [0, x_{\theta_{i0}}); \\ 1 - \frac{x_0(1-G_i(x_0^-))}{x}, & \text{if } x \in [x_{\theta_{i0}}, x_{i0}); \\ G_i(x), & \text{if } x \in [x_{i0}, 1], \end{cases}$$

By similar arguments, $\{\tilde{G}_i^{\theta_{i0}}\}_i$ generates weakly less revenue than $\{G_i\}$.

Second, define another signal distribution $\hat{G}_i^{x_{i1}}$ similar to that in Appendix A.3:

$$\hat{G}_i^{x_{i1}}(x) = \begin{cases} 1 - \frac{1-\tilde{G}_i^{\theta_{i0}}(x_{i1})}{x}, & \text{if } x \in [1 - \hat{G}_i^{\theta_{i0}}(x_{i1}), x_{i1}]; \\ \tilde{G}_i^{\theta_{i0}}(x), & \text{if } x \in (x_{i1}, 1]. \end{cases}$$

By similar arguments, $\{\hat{G}_i^{x_{i1}}\}_i$ generates strictly less revenue than $\{\tilde{G}_i^{\theta_{i0}}\}_i$. $\square$

Proof of regularity. For each buyer i , let G_i induce ironed virtual value on $[x_{i1}, x_{i2}]$. First, we construct a distribution $\hat{G}_i^{\theta_{i0}}$ similar to that in Appendix A.4:

$$\hat{G}_i^{\theta_{i0}}(x) = \begin{cases} \theta_{i0}, & \text{if } x \in [0, x_{\theta_{i0}}), \\ \hat{G}_i(x), & \text{if } x \in [x_{\theta_{i0}}, 1], \end{cases}$$

where

$$\hat{G}_i(x) = \begin{cases} G_i(x), & \text{if } x \notin [x_{i1}, x_{i2}]; \\ 1 - \frac{(1-G_i(x_{i1}))(x_{i1}-k)}{x-k}, & \text{if } x \in [x_{i1}, x_{i2}]. \end{cases}$$

By similar arguments, $\{\hat{G}_i^{\theta_{i0}}\}_i$ generates weakly less revenue than $\{G_i\}$.

Although this modification still puts a positive mass on signal 0, by the nonnegativity part of Lemma 11, we can further modify this distribution into the one with nonnegative virtual values and thereby achieve strictly less seller revenue. Thus, the seller-worst distribution must also be regular. $\square$

A.10 Proof of Proposition 1

Proof for $n = 2$. Within the class documented in the context, let θ_{0i} denote the mass on signal 0 and $1 - \theta_i$ denote the mass on signal 1 for buyer i . Without loss of generality, we assume that $\theta_{01} \leq \theta_{02}$, and then the buyer-optimal information design problem is

$$\begin{aligned} \max_{\theta_{0i}, \theta_i} \quad & \int_0^1 x d(G_1 G_2) - (1 - \theta_1 \theta_2) \\ \text{s.t.} \quad & (1 - \theta_i)(1 - \log(1 - \theta_i) + \log(1 - \theta_{0i})) = p, \quad \forall i = 1, 2. \end{aligned}$$

And the objective can be rewritten as

$$\begin{aligned} & 1 - \int_0^1 G_1 G_2 dx - (1 - \theta_1 \theta_2) \\ = & 1 - \int_0^{x_1} \theta_{01} \theta_{02} dx - \int_{x_1}^{x_2} \theta_{02} \left(1 - \frac{a_1}{x}\right) dx - \int_{x_2}^1 \left(1 - \frac{a_1}{x}\right) \left(1 - \frac{a_2}{x}\right) dx - (1 - \theta_1 \theta_2) \\ = & 1 - \theta_{01} \theta_{02} x_1 - \theta_{02} (x_2 - x_1 - a_1 (\log(x_2) - \log(x_1))) - (1 - x_2) \\ & - (a_1 + a_2) \log(x_2) - a_1 a_2 \left(1 - \frac{1}{x_2}\right) - (1 - \theta_1 \theta_2) \\ = & 2(1 - \theta_1)(\theta_{02} - \theta_2) + \theta_{02}(1 - \theta_1) (\log(1 - \theta_{01}) - \log(1 - \theta_1)) \\ & + (2 - (1 - \theta_1)\theta_{02} - \theta_1 - \theta_2) (\log(1 - \theta_{02}) - \log(1 - \theta_2)). \end{aligned}$$

Therefore, the Lagrangian is

$$\begin{aligned} \mathcal{L} = & 2(1 - \theta_1)(\theta_{02} - \theta_2) + \theta_{02}(1 - \theta_1) (\log(1 - \theta_{01}) - \log(1 - \theta_1)) \\ & + (2 - (1 - \theta_1)\theta_{02} - \theta_1 - \theta_2) (\log(1 - \theta_{02}) - \log(1 - \theta_2)) \\ & - \sum_{i=1}^2 \lambda_i ((1 - \theta_i)(1 - \log(1 - \theta_i) + \log(1 - \theta_{0i})) - p). \end{aligned}$$

Taking the first derivative with respect to θ_{01} yields

$$\frac{\partial \mathcal{L}}{\partial \theta_{01}} = \frac{-\theta_{02}(1 - \theta_1)}{1 - \theta_{01}} + \frac{\lambda_1(1 - \theta_1)}{1 - \theta_{01}} = \frac{(\lambda_1 - \theta_{02})(1 - \theta_1)}{1 - \theta_{01}}.$$

Since λ_1 is constant, $\frac{\partial \mathcal{L}}{\partial \theta_{01}}$ is either always non-positive or always nonnegative. Therefore, the optimal θ_{01} is a boundary solution. That is $\theta_{01} = \theta_{02}$ or $\theta_{01} = 0$.

First, if $\theta_{01} = \theta_{02}$, then the solution becomes the symmetric buyer-optimal information structure design problem when $n = 2$.

Second, if $\theta_{01} = 0$, then the information designer only needs to choose the optimal θ_{02} to maximize the buyers' surplus. Since $p < r_b$, the symmetric buyer-optimal information structure puts a positive mass on signal 0. Hence, $\theta_{02} > 0$; otherwise, by the mean constraint, both buyers distributions become identical and place mass only on virtual values 0 and 1, which, by Theorem 2, is not optimal.

However, to guarantee that the asymmetric case with $\theta_{01} = 0$ and $\theta_{02} > 0$ is not vacuous, we have to know the exact value of λ_1 and θ_{02} . We appeal to simulation. For instance, let $p = 0.4$, under the symmetric buyer-optimal information, the mass on signal 0 is $\theta_0 \simeq 0.1251$. Then, the mean constraint

implies $\theta_b \simeq 0.8581$. Hence, the buyers' surplus is 0.3082. If we take $\theta_{01} = 0$ and $\theta_{02} = 0.3$, then the mean constraint implies that $\theta_1 \simeq 0.8677$ and $\theta_2 \simeq 0.8374$. The buyers' surplus is $0.3107 > 0.3082$. Moreover, under the resulting asymmetric information structure, buyer 1's surplus is $0.1443 < 0.1541 = 0.3082/2$, whereas buyer 2's surplus is $0.1663 > 0.1541$. Hence, buyer 2 benefits from the asymmetric information structure more than the loss incurred by buyer 1. $\square$

Proof for $n \rightarrow \infty$. Let θ_0 denote the mass on signal 0 and $1 - \theta$ denote the mass on signal 1, let buyer i 's signal distribution be

$$G_i(x) = \begin{cases} \theta_0, & \text{if } x \in [0, x_1); \\ 1 - \frac{(1-\theta)(1-p)}{x-p}, & \text{if } x \in [x_1, 1); \\ 1, & \text{if } x = 1, \end{cases}$$

where $x_1 = p + (1 - \theta)(1 - p)/(1 - \theta_0)$, and all the other buyers' signal distributions be the degenerate distribution as in Corollary 1. First, for $x \in [x_1, 1)$, the induced virtual value is p . And since $x_1 > p$, by the allocation rule, the good will be allocated to buyer i if buyer i 's signal belongs to $[x_1, 1)$. Since the seller can only get p which is strictly less than x_1 when $x_i \in [x_1, 1)$, buyer i must obtain some positive information rents. Hence, the buyers' total surplus is strictly positive.²⁷ $\square$

A.11 Proof of Proposition 2

Proof. With change of variable, for $n = 2$, the seller-worst information designer's problem can be written as

$$\begin{aligned} \max_{\{F_i(k)\}_{i=1}^2} & \int_0^1 \prod_{i=1}^2 F_i(k) dk - 1 \\ \text{s.t.} & \int_0^1 (1 - F_i(k))(1 - \log(1 - F_i(k))) dk = p_i, \quad \forall i = 1, 2. \end{aligned} \quad (33)$$

Consider the following Lagrangian $\mathcal{L}$:

$$\mathcal{L}(F_i(k), \lambda_i) = \int_0^1 \prod_{i=1}^2 F_i(k) - \sum_{i=1}^2 \lambda_i ((1 - F_i(k))(1 - \log(1 - F_i(k)))) dk + \sum_{i=1}^2 \lambda_i p_i.$$

By Theorem 4.2.1 of [van Brunt \(2004\)](#), for any state k , the Euler-Lagrange equation with respect to $F_i(k)$ is

$$\prod_{j \neq i} F_j(k) - \lambda_i \log(1 - F_i(k)) = 0, \quad \forall i = 1, 2.$$

Denote $\theta_i = F_i(k)$ and we have

$$\theta_2 = \lambda_1 \log(1 - \theta_1),$$

$$\theta_1 = \lambda_2 \log(1 - \theta_2).$$

²⁷For example, let $p = 0.4$, take $\theta_0 = 0.4751$ and $\theta = 0.8661$ which satisfies the mean constraint. Then, the buyers' surplus is 0.1097 which is strictly larger than 0 under the symmetric buyer-optimal information structure when $n \rightarrow \infty$.

First, note that λ_i should be negative. Then, we claim that there is a unique solution pair (θ_1^*, θ_2^*) such that the Euler-Lagrange equations are satisfied. Hence, given any solution (θ_i, θ_j) , we have

$$\theta_1 = \lambda_2 \log(1 - \lambda_1 \log(1 - \theta_1)). \quad (34)$$

Taking the first and second derivative of the right-hand side, we obtain

$$\begin{aligned} \frac{\partial}{\partial \theta_1} \lambda_2 \log(1 - \lambda_1 \log(1 - \theta_1)) &= \frac{\lambda_1 \lambda_2}{(1 - \theta_1)(1 - \lambda_1 \log(1 - \theta_1))} > 0, \\ \frac{\partial^2}{\partial \theta_1^2} \lambda_2 \log(1 - \lambda_1 \log(1 - \theta_1)) &= \frac{\lambda_1 \lambda_2 (1 - \lambda_1 \log(1 - \theta_1) - \lambda_1)}{(1 - \theta_1)^2 (1 - \lambda_1 \log(1 - \theta_1))^2} > 0. \end{aligned}$$

Hence, the right-hand side is convex and increasing in θ_1 . It follows that there will be only one solution of θ_1 such that Equation (34) holds.

Hence, an optimal $F_1^*(k)$ and $F_2^*(k)$ are both constantly equal to θ_1 and θ_2 for $k < 1$. That is, both $F_1^*(k)$ and $F_2^*(k)$ have binary support $\{k, 1\}$. By part (1) of Lemma 10, the uniqueness of θ_i implies (F_1^*, F_2^*) with the binary support is also a global maximizer.

Again, let θ_i be the mass on the virtual value k for buyer i . Then the information design problem is reduced into

$$\begin{aligned} \max_{k, \theta_1, \theta_2} & (1 - k)\theta_1\theta_2 - 1 \\ \text{s.t. } & k + (1 - k)((1 - \theta_i)(1 - \log(1 - \theta_i))) = p_i. \end{aligned}$$

The Lagrangian with multiplier λ_i is

$$\mathcal{L}(k, \theta_1, \theta_2, \lambda_1, \lambda_2) = (1 - k)\theta_1\theta_2 - 1 - \sum_{i=1}^2 \lambda_i (k + (1 - k)((1 - \theta_i)(1 - \log(1 - \theta_i))) - p_i).$$

The Euler-Lagrange equation with respect to θ_i is

$$\frac{\partial \mathcal{L}}{\partial \theta_i} = (1 - k)(\theta_{-i} - \lambda_i \log(1 - \theta_i)) = 0 \implies \lambda_i = \frac{\theta_{-i}}{\log(1 - \theta_i)}. \quad (35)$$

Also the Euler-Lagrange equation with respect to k is

$$\frac{\partial \mathcal{L}}{\partial k} = -\theta_1\theta_2 - \sum_{i=1}^2 \lambda_i (\theta_i + (1 - \theta_i) \log(1 - \theta_i)). \quad (36)$$

Plugging Equation (35) into Expression (36), we have

$$\frac{\partial \mathcal{L}}{\partial k} = \theta_1\theta_2 - \theta_1 \left(1 + \frac{\theta_2}{\log(1 - \theta_2)}\right) - \theta_2 \left(1 + \frac{\theta_1}{\log(1 - \theta_1)}\right).$$

We claim that $\frac{\partial \mathcal{L}}{\partial k} \leq 0$. To see this, it suffices to show that $1 + \frac{\theta_i}{\log(1 - \theta_i)} \geq \frac{\theta_i}{2}$. Indeed,

$$\frac{\partial}{\partial \theta_i} \left(- \left(1 - \frac{\theta_i}{2}\right) \log(1 - \theta_i) - \theta_i \right) = \frac{1}{2} \left(\frac{\theta_i}{1 - \theta_i} + \log(1 - \theta_i) \right) \geq 0.$$

Hence, $-\left(1 - \frac{\theta_i}{2}\right) \log(1 - \theta_i) - \theta_i \geq 0$. That is, $1 + \frac{\theta_i}{\log(1 - \theta_i)} \geq \frac{\theta_i}{2}$. Therefore, in order to maximize the objective, the information designer should choose $k = 0$. $\square$

A.12 Auctions under an irregular distribution

Suppose that the support of the signal x is $[1, 2]$. Consider the signal distribution:

$$G(x) = \begin{cases} 2x - 2, & \text{if } x \in [1, \frac{4}{3}); \\ \frac{x}{2}, & \text{if } x \in [\frac{4}{3}, 2]. \end{cases}$$

The quantile function $x(\tau)$ of distribution G is given by:

$$x(\tau) = \begin{cases} \frac{\tau}{2} + 1, & \text{if } \tau \in [0, \frac{2}{3}); \\ 2\tau, & \text{if } \tau \in [\frac{2}{3}, 1]. \end{cases}$$

The virtual value without ironing is given by

$$\varphi(x) = x - \frac{1 - G(x)}{g(x)} = \begin{cases} x - \frac{1 - (2x - 2)}{2} = 2x - \frac{3}{2}, & \text{if } x \in [1, \frac{4}{3}); \\ x - \frac{1 - x/2}{1/2} = 2x - 2, & \text{if } x \in [\frac{4}{3}, 2]. \end{cases}$$

Denote $\Phi(x) = \int_1^x \varphi(t)g(t)dt$. Then,

$$\Phi(x) = \begin{cases} \int_1^x 2(2t - 3/2)dt = (x - 1)(2x - 1), & \text{if } x \in [1, \frac{4}{3}); \\ \int_1^{4/3} 2(2t - 3/2)dt + \int_{4/3}^x \frac{1}{2}(2t - 2)dt = \frac{1}{2}x(x - 2) + 1, & \text{if } x \in [\frac{4}{3}, 2]. \end{cases}$$

Therefore, $\Phi(\tau) = \Phi(x(\tau))$ is given by

$$\Phi(\tau) = \begin{cases} (x(\tau) - 1)(2x(\tau) - 1) = \frac{\tau}{2}(\tau + 1), & \text{if } \tau \in [0, \frac{2}{3}); \\ \frac{1}{2}x(\tau)(x(\tau) - 2) + 1 = 1 - 2\tau + 2\tau^2, & \text{if } \tau \in [\frac{2}{3}, 1]. \end{cases}$$

Denote $\Psi(\tau)$ be the largest convex function such that $\Psi(\tau) \leq \Phi(\tau)$. Then

$$\Psi(\tau) = \begin{cases} \frac{\tau}{2}(\tau + 1), & \text{if } \tau \in [0, \frac{1}{2}); \\ \tau - \frac{1}{8}, & \text{if } \tau \in [\frac{1}{2}, \frac{3}{4}); \\ 1 - 2\tau + 2\tau^2, & \text{if } \tau \in [\frac{3}{4}, 1]. \end{cases}$$

Therefore, the ironed virtual value $\hat{\varphi}(\tau) = \Psi'(\tau)$ is given by

$$\hat{\varphi}(\tau) = \begin{cases} \tau + \frac{1}{2}, & \text{if } \tau \in [0, \frac{1}{2}); \\ 1, & \text{if } \tau \in [\frac{1}{2}, \frac{3}{4}); \\ 4\tau - 2, & \text{if } \tau \in [\frac{3}{4}, 1]. \end{cases}$$

Replace τ by $G(x)$ and the ironed virtual value $\hat{\varphi}(x)$ in terms of x is given by

$$\hat{\varphi}(x) = \begin{cases} 2x - \frac{3}{2}, & \text{if } x \in [1, \frac{5}{4}); \\ 1, & \text{if } x \in [\frac{5}{4}, \frac{3}{2}); \\ 2x - 2, & \text{if } x \in [\frac{3}{2}, 2]. \end{cases}$$

Note that $\hat{\varphi}(x) \geq \hat{\varphi}(1) = 1/2$ is always positive. Hence, the optimal reserve price is zero. Let us then compute the expected highest ironed virtual value $\hat{\varphi}(x)$ and the revenue under a second-price auction.

1. First, by symmetry, the largest value $x^{(1)}$ induces the highest ironed virtual value and the highest value $x^{(1)}$ follows the distribution G^2 . Then, we have

$$\begin{aligned}\mathbb{E}[\hat{\varphi}(x^{(1)})|G] &= \int_1^2 \hat{\varphi}(x) dG^2(x) = \int_1^2 \hat{\varphi}(x) 2 \cdot g(x) G(x) dx \\ &= \int_1^{5/4} (2x - 3/2) 2 \cdot 2(2x - 2) dx + \int_{5/4}^{4/3} 1 \cdot 2 \cdot 2(2x - 2) dx + \int_{4/3}^{3/2} 1 \cdot \frac{2x}{4} dx + \int_{3/2}^2 (2x - 2) \cdot \frac{2x}{4} dx \\ &= \frac{5}{24} + \frac{7}{36} + \frac{17}{144} + \frac{2}{3} = \frac{19}{16}.\end{aligned}$$

2. Second, the lowest value $x^{(2)}$ follows the distribution $2G - G^2$, then we have

$$\begin{aligned}\mathbb{E}[x^{(2)}|G] &= \int_1^2 x d(2G - G^2) = \int_1^2 x \cdot 2 \cdot g(x)(1 - G(x)) dx \\ &= \int_1^{4/3} x \cdot 2 \cdot 2(3 - 2x) dx + \int_{4/3}^2 x \cdot 2 \cdot \frac{1}{2} (1 - \frac{x}{2}) dx \\ &= \frac{82}{81} + \frac{14}{81} = \frac{96}{81} = \frac{32}{27} = \frac{19}{16} - \frac{1}{432} < \frac{19}{16}.\end{aligned}$$

Therefore, a second-price auction with an optimal reserve price 0 obtains strictly less revenue than the expected highest ironed virtual value. Hence, the second-price auction with an optimal reserve price 0 is not an optimal auction.

Note that although $G(x)$ can induce the ironed virtual value $\hat{\varphi}(x)$, the regular distribution $\hat{G}$ which also induces the same virtual value $\hat{\varphi}(x)$ will first-order stochastically dominate G . Hence, the expectation of the lowest/second-highest value $x^{(2)}$ of $\hat{G}$ is strictly larger than that of G . Formally, $\hat{G}$ is given by

$$\hat{G}(x) = \begin{cases} 2x - 2, & \text{if } x \in [1, \frac{5}{4}]; \\ 1 - \frac{1}{8(x-1)}, & \text{if } x \in (\frac{5}{4}, \frac{3}{2}); \\ \frac{x}{2}, & \text{if } x \in [\frac{3}{2}, 2]. \end{cases}$$

For $x \in [1, \frac{5}{4}] \cup [\frac{3}{2}, 2]$, we have $\hat{G}(x) = G(x)$. For $x \in (\frac{5}{4}, \frac{3}{2})$, $\hat{G}(x) < G(x)$. Hence $\hat{G}$ first-order stochastically dominates G . Moreover, we have $\mathbb{E}[x^{(2)}|\hat{G}] = \frac{19}{16} = \mathbb{E}[\hat{\varphi}(x^{(1)})|G]$. Of course, this is an example of the standard result that the optimal auction is a second-price auction with an optimal reserve price, provided that the signal distribution is regular.

A.13 Inequivalence under a continuous prior with negative virtual values

Claim 2. The buyer-optimal information structure and the seller-worst information structure are inequivalent under the following regular continuous prior $\hat{H}$ with $\mathbb{E}_{\hat{H}}[x] \simeq 0.26785$,

$$\hat{H}(x) = \begin{cases} 5x, & \text{if } x \in [0, 1/20]; \\ 1 - \frac{9/80}{x - (-1/10)}, & \text{if } x \in [1/20, 0.999]; \\ \frac{112500}{1099}x - \frac{111401}{1099}, & \text{if } x \in [0.999, 1]. \end{cases}$$

It follows that $\hat{H}$ induces negative virtual values on $[0, 0.999]$. In particular, the virtual value is $-1/10$ on $[1/20, 0.999]$.

Proof. First, we provide an upper bound of the buyers' surplus under some seller-worst information structure G_s^* .

- Suppose that we only have the mean constraint instead of the mean-preserving spread constraint. By Theorem 1, the following signal distribution is the unique seller-worst information structure,

$$G_s = \begin{cases} 1 - \frac{x_s}{x}, & \text{if } x \in [x_s, 1]; \\ 1, & \text{if } x = 1, \end{cases}$$

where $x_s \simeq 0.07445$ solves $x_s(1 - \log(x_s)) = \mathbb{E}_H[x]$. Since the information designer faces a more strict mean-preserving spread constraint under the prior H , the seller-worst information structure G_s^* will generate a weakly higher seller's revenue than G_s does.

- Even though G_s is infeasible under the mean-preserving spread constraint, we claim that the seller-worst information structure G_s^* must generate a lower buyers' surplus than G_s does.

Since the seller-worst signal distribution must be regular and admit only nonnegative virtual values even under continuous priors, G_s must be a mean-preserving spread of G_s^* .

- To explain, since the virtual value is $\varphi(x) = x - \frac{1-G(x)}{G'(x)}$, at an intersection of two regular distributions, both of the distributions have the same x and $G(x)$, and thereby the distribution with a higher virtual value $\varphi(x)$ must have a higher slope $G'(x)$ at $(x, G(x))$. Moreover, G_s has zero virtual values on $[x_s, 1)$ and G_s has nonnegative virtual values on $[0, 1]$. As illustrated in Figure 6, we claim that G_s^* will only cross G_s from the below. Suppose G_s^* crosses G_s from the above at x_0 as the blue curve in Figure 6. Since G_s has a higher slope than G_s^* at $(x_0, G_s(x_0))$, $\varphi(x_0|G_s^*) < \varphi(x_0|G_s) = 0$ which contradicts with the fact that G_s^* must induce nonnegative virtual values. Therefore, G_s^* must cross G_s from the below once and only once.

Therefore, by Lemma 2, G_s generates more total surplus than G_s^* . Since the good is always allocated under both G_s^* and G_s , the buyers' surplus (as the total surplus minus the seller's revenue) under G_s^* will be lower than that under G_s .

- The buyers' surplus under G_s is 0.24898, hence the buyers' surplus under the seller-worst signal distribution G_s^* can not exceed 0.24898.

Second, to show the inequivalence, it suffices to find another feasible signal distribution $\hat{G}$ which generates a higher buyers' surplus than 0.24898. And the construction is as follows,

$$\hat{G}(x) = \begin{cases} 5x, & \text{if } x \in [0, 1/50]; \\ 0.1, & \text{if } x \in (1/50, x_0); \\ 1 - \frac{0.9x_0}{x}, & \text{if } x \in [x_0, 0.999]; \\ 1, & \text{if } x \in [0.999, 1], \end{cases}$$

where $x_0 \simeq 0.0858$.

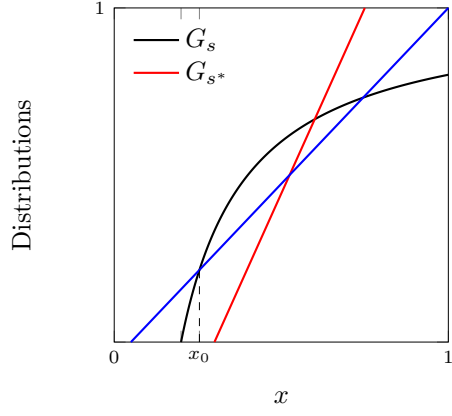


Figure 6: G_{s^*} single-crosses G_s

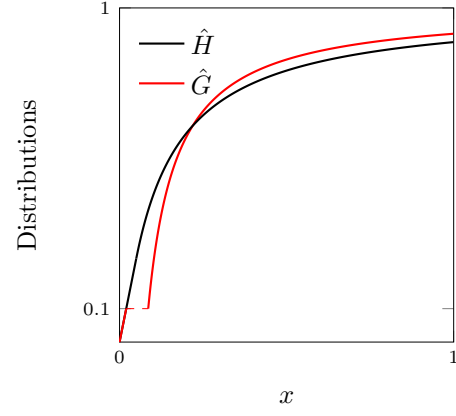


Figure 7: $\hat{G}$ is feasible

The yield buyers' surplus under $\hat{G}$ is $0.25205 > 0.24898$.

Finally, we use Figure 7 to illustrate that $\hat{H}$ is a mean-preserving spread of $\hat{G}$. □