How Competition Shapes Information in Auctions*

Agathe Pernoud[†] Simon Gleyze[‡]

August 2024

Abstract

We consider auctions where buyers can acquire costly information about their valuations and those of others, and investigate how competition between buyers shapes their learning incentives. In equilibrium, buyers find it cost-efficient to acquire some information about their competitors so as to only learn their valuations when they have a fair chance of winning. We show that such learning incentives make competition between buyers less effective: losing buyers often fail to learn their valuations precisely and, as a result, compete less aggressively for the good. This depresses revenue, which remains bounded away from what the standard model with exogenous information predicts, even when information costs are negligible. Finally, we highlight some implications for auction design.

*We are grateful to Matthew Jackson and Paul Milgrom for invaluable guidance and support. We also thank Mohammad Akbarpour, Nina Bobkova, Ben Brooks, Eric Budish, Piotr Dworczak, Liran Einav, Ellen Muir, Doron Ravid, Ilya Segal, Andy Skrzypacz for helpful discussions as well as seminar participants at Stanford and other universities for their questions and comments. This research was supported by the E.S. Shaw and B.F. Haley Fellowship for Economics through a grant to the Stanford Institute for Economic Policy Research.

[†]Booth School of Business, University of Chicago. Email: agathe.pernoud@chicagobooth.edu

[‡]Uber. Email: gleyze.simon@gmail.com

1 INTRODUCTION

In many auctions, participants spend significant time and resources learning about the goods for sale before bidding. Relevant examples are the sales of complex, high-value assets such as companies, broadband licenses, or procurement contracts, during which interested buyers conduct thorough due diligence. For instance, bidders in takeover auctions get access to extensive information about the target's operations and finances, allowing them to assess synergies and estimate how much they value its acquisition.

Accessing and processing such information is often costly, and buyers only want to undertake this investment if they have a decent chance of winning the auction.[1] Effective auction design then needs to account for buyers' learning incentives, and existing literature has mostly focused on the *intensive* margin—that is, *how much* information buyers acquire about their values under various auction formats (see, e.g., Persico (2000); Bergemann and Välimäki (2002); Shi (2012)). Much less attention has been paid to how auction design affects *what types* of information they seek.

In practice, auction participants have access to a wide range of information and have ample flexibility as to what to learn about. Some information helps them assess their own willingness to pay for the good while other pertains to the rival bids they will face. Anecdotal evidence suggests that participants in high-stake auctions seek both types of information. For instance, GTE's bidding team prepared for the 1994 spectrum auction run by the Federal Communications Commission by assessing both the value of each license to GTE and the level of competition in each market (Salant (1997)). To assess competition, they looked at which bidders were eligible to bid on which licences, estimated their budgets, and identified synergies between each license and their rivals' existing networks.[2] This gave GTE a better sense of which licences they could reasonably win, which were then the focus of bid preparation efforts.

Naturally, what buyers choose to learn is likely to affect how they bid, and thus the outcome of the auction. In the above-mentioned FCC auction, the Los Angeles license raised "little" revenue,[3] perhaps because the participation of Pacific Telesis, the

[1]Due diligence for the acquisition of a company can take months and involves legal and accounting fees that average to 47 basis points of the deal value (Cole et al. (2016)), in addition to consulting fees.

[2]Buyers have access to some information about their competitors in other settings too. For instance, bidders in highway procurement auctions can learn about the location of their rivals' machines and their backlogs, and bidders in takeover auctions about their rivals' financial constraints.

[3]It was sold at the price of $26 per capita, which was considered low as less profitable licenses were sold at higher prices, e.g., the Chicago license was sold at $31 per capita. See

main company in California, scared away the competitors. GTE's bidding team indeed seemed to think they had little chance of winning that license, and so focused their efforts elsewhere (Salant (1997)). Beyond such anecdotal evidence, we do not however have a good theoretical understanding of what information bidders seek when given the flexibility, nor of how it affects the performance of the auction.

This paper proposes a tractable model of multidimensional learning in second-price auctions. We focus on the canonical auction setting with independent private values, in which buyers' valuations are drawn i.i.d. from some common knowledge distribution. Importantly, buyers do not know their valuations ex-ante and can acquire costly information before bidding. The main innovation of our model is that it gives buyers flexibility in *what* information they can seek. Specifically, buyers can acquire two signals—one about their own valuations for the good and one about those of their competitors—and can choose in which order to acquire them. They also have some flexibility in choosing each signal's informativeness and, in particular, can choose how each signal partitions the set of possible valuations. Information is costly, and we require that the cost satisfies an appropriate notion of convexity.

Our first main result is that buyers do not converge to becoming fully informed of their valuations in equilibrium, even as information costs become arbitrarily small (Theorem 1). Instead, they find it cost-efficient to first assess the valuation of their toughest competitor, and only then learn about their own, which they do only when they have a chance of winning. As a result, their private information when entering the auction (i.e., their types) are interdependent: not only do buyers have information that is relevant to others, but their own expected valuations may depend on what they learned about the competitors. We characterize equilibrium information structures in high-stake auctions (i.e., auctions where information costs are small relative to the value of the good) and show that buyers only learn their valuations if it falls in a similar range as that of their toughest competitor. The information buyers acquire is then deeply shaped by the competitive pressure they impose on each other.

We then examine how buyers' learning incentives, in turn, affect the performance of the second-price auction. We show that expected revenue remains bounded away from what the standard model predicts, even as information costs vanish (Theorem 2). Indeed, losing buyers often fail to learn their valuations precisely, and since they bid

https://www.fcc.gov/auctions-summary for the auction outcome and Klemperer (2002) for a more detailed account.

their expected valuations for the good in equilibrium, this leads to a regression to the mean of bids. Losing bids are then less dispersed than in the standard model, which depresses the expected second-highest bid and hence expected revenue whenever the number of buyers $N \geq 3$.[4]

Our main results highlight a new adverse effect of competition on revenue, and we investigate some implications for auction design. We show that, whenever the number of bidders N is not too small, attracting an additional bidder is less valuable than setting an optimal reserve price, suggesting that the seminal result of Bulow and Klemperer (1996) relies on buyers knowing their valuations fully. There are several forces at play. First, an additional bidder does not raise revenue as much as in the standard model, as that bidder does not learn his value fully. Second, a carefully chosen reserve price is more valuable as losing bidders oftentimes fail to learn their valuations for the good, leaving a larger expected gap between the highest and second-highest bids.

We then show how the seller can mitigate the revenue loss by randomizing access to the auction. Indeed, by only allowing a randomly-chosen subset of buyers to participate in the sale, the seller reduces buyers' incentives to learn about their competitors: even if a strong buyer is present, he might not be granted access; others then still have a chance of winning, and hence an incentive to learn about their own valuations. In high-stake auctions, this unambiguously improves expected revenue. This result might explain why bidders in takeover auctions are often required to sign non-disclosure agreements preventing them from revealing, among other things, their participation in the sale. Such agreements are all the more important as buyers' incentives very much conflict with the seller's on that point: high-valuation buyers benefit from disclosing their participation and bids to others, so as to deter them from conducting due diligence and reduce the expected price.[5]

Finally, we close the paper with a discussion of the model's key assumptions, which are the ones imposed on the process of information acquisition. We see Theorems 1 and 2 as the main results of the paper, and show that they are robust to a weakening of these assumptions and alternative model specifications. The market design implications rely more heavily on our assumptions, but capture practically relevant insights.

[4]This effect persists as information costs become arbitrarily small, thus highlighting a discontinuity between the standard model where buyers know their valuations ex-ante (information is then effectively free), and ours.

[5]Even if buyers are legally bound not to disclose their bidding intentions, they might still be able to signal them—e.g., toeholds can sometimes signal an intention to bid aggressively in takeover auctions.

1.1 Related Literature

First, we build on a previous paper (Gleyze and Pernoud (2022)), which investigates whether the seller can design a mechanism under which participants only have an incentive to learn about their own preferences. We show that it is impossible: most selling mechanisms incentive participants to learn about others' preferences as well, leading their types to be endogenously interdependent. Gleyze and Pernoud (2022) however does not characterize what an equilibrium information structure looks like—it just shows that it has to feature some interdependencies, but not what form they take—nor how it affects the performance of the selling mechanism. The present paper addresses these questions in the context of the second-price auction.[6]

Second, our paper contributes to the literature on entry and learning costs in auctions. Levin and Smith (1994) characterize the symmetric equilibrium when buyers pay a fixed cost to learn their values before bidding in a second-price auction. They show that equilibrium entry decisions are revenue-maximizing in the IPV setting.[7] Several papers highlight the benefits of dynamic auction formats (Compte and Jehiel (2007)), and in particular of the Dutch auction (Miettinen (2013); Kleinberg et al. (2016)), in coordinating learning across buyers, as dynamic formats endogenously reveal information on the toughness of competition. Another strand of the literature allows buyers to flexibly choose *how much* information to acquire (Hausch and Li (1993); Persico (2000); Bergemann and Välimäki (2002); Shi (2012); Kim and Koh (2022)). Importantly, these papers consider information acquisition about a one-dimensional random variable—usually buyers' own valuations or a component that is common to all buyers.

Our key contribution is to propose a model of *multidimensional* learning in which buyers can separately choose how much to learn about self and others. To our knowledge, the only papers that study multidimensional learning in auctions are Larson and Sandholm (2001b,a) and Bobkova (2019). Larson and Sandholm (2001b,a) also consider learning about both self and others. They show that computationally limited agents have an incentive to learn about others in Vickrey and ascending auctions and

[6]We focus on the second-price auction, not only because it is a widely-used auction format but also because it is *strategy-proof*: if buyers knew their valuations, they would have a dominant strategy and would have no incentive to inquire about the competition. We can thus isolate the detrimental effect of competition on learning incentives.

[7]A related literature assumes that buyers have some private information when making entry decisions (Ye (2007); Quint and Hendricks (2018); Lu et al. (2021), etc.). Entry then serves as a screening mechanism that the seller can leverage. Such considerations are absent in our paper.

compare different models of costly deliberation. In Bobkova (2019), buyers' valuations are composed of a private and a common component, and buyers choose how much to learn about each.[8] She shows that buyers only seek information about their private components in second-price auctions, which contrasts with our results.

A relatively small literature studies buyers' incentives to learn about the competition in first-price auctions (Tian and Xiao (2007)) and auctions with interdependent values (Kim and Koh (2020)). Information about opponents' types is valuable as it allows buyers to either shade their bids more aggressively (in the former) or alleviate the winner's curse (in the latter). Such incentives are absent in our setting as buyers compete in a second-price auction and their valuations are independent and private.

Our paper also speaks to the literature highlighting the value of competition in selling mechanisms. Bulow and Klemperer (1996) show that attracting just one more buyer has more value than using an optimal reserve price. Relatedly, Bulow and Klemperer (2009) show that with costly entry, actual competition in an auction dominates potential competition from a sequential entry mechanism.[9] These results, however, take buyers' information as fixed. Instead, our paper asks how competition affects the information buyers acquire and reaches opposite conclusions. Gershkov et al. (2021) also qualify the value of competition and show that it can even hurt revenue if buyers can invest to increase their values before bidding.

Finally, several papers study the performance of auctions when buyers are ex-ante asymmetric (Maskin and Riley (2000); Compte and Jehiel (2002); Kim and Che (2004); Cantillon (2008); Jehiel and Lamy (2015); Marquez (2021)). In particular, Marquez (2021) shows that the presence of an asymmetrically strong buyer can hurt revenue. In our paper, buyers are ex-ante identical, but we show that asymmetries in private information arise endogenously, even in symmetric equilibria.

We end by noting that an extensive literature examines the different, though related, question of optimal information *disclosure* in auctions (see, e.g., Bergemann and Pesendorfer (2007); Prummer and Nava (2023) and references therein). In our paper, the seller has no information and does not intervene in buyers' learning process.[10]

[8] They cannot however learn about their competitors' values directly.

[9] Roberts and Sweeting (2013) extend their model with ex-ante noisy signals and asymmetries and find that the sequential entry mechanism can dominate the auction.

[10] Note that the equilibrium distribution of types in our setting is highly subopimal in the eyes of the seller since it leads to a revenue loss compared to full information.

1.2 A Motivating Example

Two buyers compete in a second-price auction to acquire a good. Their values are drawn i.i.d. from a finite set of possible valuations V. For the sake of this example, assume that buyer j is exogenously informed of his valuation v_j, and plays his dominant strategy, such that he always truthfully bids his valuation. For simplicity, the example focuses on the incentives of buyer i.

Buyer i does not know his valuation ex-ante but can acquire costly information before bidding. Importantly, buyer i has flexibility as to what to learn about: he can acquire a signal about his own valuation v_i and one about his opponent's v_j. He has full flexibility in how to design each of these two signals and can furthermore choose in which order to acquire these signals. Information is costly and, for this example only, we consider the entropic cost function. Informally, the cost of each signal is proportional to the expected reduction in uncertainty as measured by the entropy of beliefs:

$$\text{cost of a signal} = \lambda\Big(\text{prior entropy} - \mathbb{E}[\text{posterior entropy}]\Big)$$

where λ is a scaling parameter.

We examine what information buyer i acquires as the size of information costs λ goes to zero. From buyer i's perspective, this is simply a decision problem: the behavior of buyer j is fixed, so it is as if i faced a random price v_j and could learn both about the realized price and his realized valuation for the good. For small enough information costs, buyer i must learn enough to avoid (non-trivial) mistakes. That is, i must learn enough so that he (almost always) submits a winning bid $b_i > v_j$ whenever $v_i > v_j$ and a losing bid $b_i < v_j$ whenever $v_i < v_j$.[11]

Interestingly, many learning strategies guarantee that buyer i makes no mistakes when bidding. For instance, buyer i can learn his valuation fully and learn nothing about his competitor's. That is, i can partition the state space $V \times V$ as depicted in Figure 1 (left panel) and bid his value v_i in any state (v_i, v_j). Another such learning strategy is depicted in the right panel of Figure 1: i first learns whether the opponent's value is above or below some threshold v^*; if $v_j > v^*$ (resp. $v_j \leq v^*$), he learns his value fully if it is above v^* (resp. below v^*) but bundles together all values below v^* (resp. above v^*). This is enough for i to bid optimally in the auction as well, as he learns

[11] Let us ignore ties for the sake of this example. The main model will add some noise to buyers' valuations to break ties.

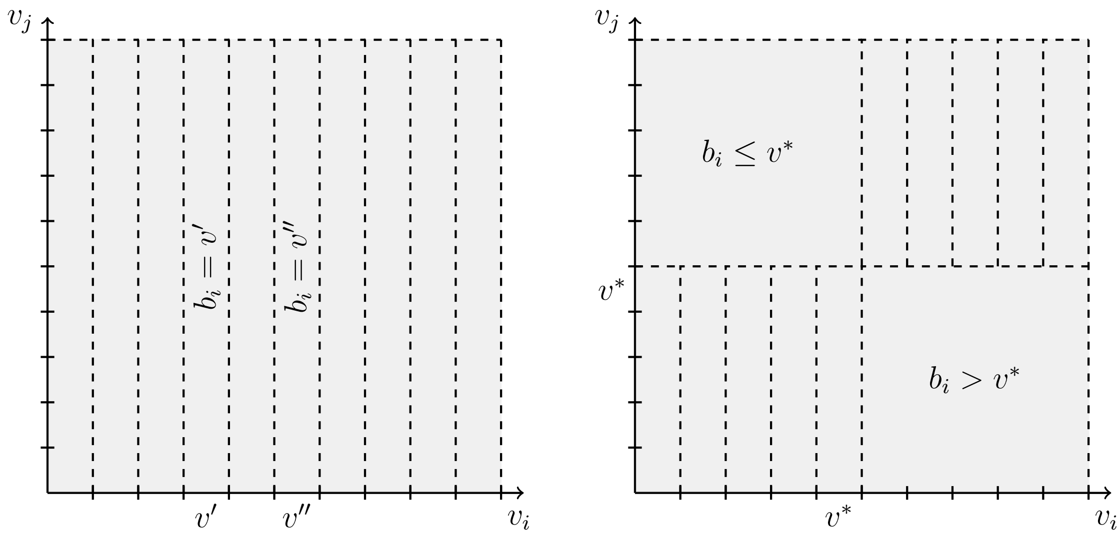


Figure 1: Two learning strategies that guarantee buyer i bids optimally, leading to two different partitions of the state space $V \times V$.

enough to know whether or not it is optimal for him to win.

As information costs vanish, buyer i optimally chooses the *cheapest* information structure that ensures he makes no mistakes in the auction. In this paper, we show that when information costs are sufficiently convex in a signal's informativeness, fully learning one's value is not cost-efficient. For instance, when there are $|V| = 10$ possible valuations and they are all ex-ante equally likely, learning v_i fully costs $\lambda \log 10$ under the entropic cost. The information structure depicted in the right panel of Figure 1 however only costs $\lambda[\log 2 + \log 10 - 0.5 \log 5]$, which is strictly cheaper.

This example assumes that buyer j knows and bids his value to simplify the argument, but the same reasoning works in equilibrium when all buyers simultaneously choose how to learn and bid. The added difficulty is that if i fails to learn his valuation *fully* and j learns something about v_i, then j has information that is relevant to i. Buyers' information—or *types*—is then interdependent and characterizing equilibrium bidding is much less straightforward. In our model, we thus add structure to the learning process and require signals to be monotone partitions of V.[12] This allows us to characterize the equilibrium information structure as information costs vanish, which looks like the one in Figure 1 (right).

[12]Note that, in this motivating example, a monotone partition is the optimal type of signal that i would acquire about his value as information costs vanish.

2 The Model

We introduce a tractable model of multi-dimensional learning in auctions. A seller puts a unique, indivisible good for sale through a second-price auction.[13] There are N buyers, and buyer i's valuation for the good is denoted by ν_i. A buyer's valuation is the sum of two components $\nu_i = v_i + u_i$, where $v_i \in V$ should be interpreted as the main component—we sometimes abuse language and refer to v_i as a buyer's valuation—and $u_i \in U$ as small mean-zero noise. Both components are identically and independently distributed across buyers. Main components $(v_i)_i$ are drawn i.i.d. from a finite set $V \subset \mathbb{R}_+$ according to a probability distribution $p \in \Delta V$. Noise terms $(u_i)_i$ are drawn from a compact interval $U \equiv [\underline{u}, \overline{u}] \subset \mathbb{R}$ according to a strictly positive and continuous density, with $\mathbb{E}[u_i] = 0$. They are small, in the sense that $\min_{v_i' \neq v_i''} |v_i' - v_i''| > \overline{u} - \underline{u}$. Hence if a buyer has a strictly greater v_i than another, then he must necessarily have a strictly greater overall valuation ν_i. The noise terms are only included in the model to address technical issues arising from the discreteness of V, but serve no other purpose.[14] We take them to be sufficiently small so as not to interfere with the rest of our analysis.

Buyers have quasilinear utility functions. Buyer i's gross payoff from the auction in state $(\nu_j)_j$ at bid profile $(b_j)_j$ equals

$$U(\nu_i, b_i, b_{-i}) \equiv \begin{cases} \dfrac{\nu_i - \max_{j \neq i} b_j}{|\{j = 1, \dots N \text{ s.t. } b_j = b_i\}|} & \text{if } b_i = \max\limits_j b_j \\ 0 & \text{otherwise} \end{cases}.$$

Note that we are considering a setting in which buyers' valuations are independent and private. Hence if buyers knew their own valuations ν_i, it would be a dominant strategy for them to bid truthfully in the auction, and the seller's expected revenue would be the expected second-highest value.

[13] The intuition underlying our results seems more general and should extend to other environments—e.g., R&D races and other types of contests, etc.

[14] To understand how the noise terms matter, consider what happens when there are none, and buyers fully learn their valuations. Then equilibrium bids take values in V, and no bid is ever made *strictly in between* two neighboring values. A buyer has then no incentive to learn to distinguish such values. In equilibrium, buyers would have to randomize between bundling and not bundling neighboring values. The noise terms can be interpreted as perturbations à la Harsanyi (1973): they allow us to dispense from such randomization and guarantee that buyers cannot predict others' preferences *perfectly*.

Information Structures. Buyers start with no private information, but they can learn about the realization of $(\tilde{v}_i)_i$ at some cost before competing in the auction. They also learn their own (and only their own) noise terms $\tilde{u}_i$ for free at the end of the information acquisition process.

We assume that buyers can acquire two signals, one about their own valuations $\tilde{v}_i$ and one about others' $\tilde{v}_{-i}$. Without loss of optimality, buyers first acquire information about others' valuations and, conditional on the realization of this signal, acquire information on their own. This is without loss in the sense that all our results go through if we allow buyers to choose in which order to acquire these two signals (see Section 6 for a formal argument). To reduce the dimension of the problem, we furthermore assume that buyers can only learn about $\max_{j\neq i}\tilde{v}_j$, and not the full vector $\tilde{v}_{-i}$.[15]

We model information acquisition about any random variable as the choice of a partition $\Pi = \{\pi_1, \dots \pi_L\}$ of the set of possible realizations V. That is, if a buyer chooses information partition Π, then the buyer learns to which element of the partition the realization of the random variable belongs. If the chosen partition is $\Pi = \{V\}$, then no information is acquired. If $\Pi = \{\{v\}_{v\in V}\}$, then the partition is fully revealing. We furthermore require that buyers choose *monotone* partitions, meaning that if $v', v'' \in \pi_l$ with $v' < v''$, then all $v \in (v', v'')$ also belong to the element π_l of the partition.

Information is costly. Letting $\mathcal{P}$ denote the set of all possible monotone partitions of V, the cost of a signal $c : \mathcal{P} \times \Delta(V) \longrightarrow \mathbb{R}_+$ is a function of both the chosen partition and the prior belief about the random variable of interest. Indeed, even though both $\tilde{v}_i$ and $\max_{j\neq i}\tilde{v}_j$ take realizations in the same set V, they have different prior probabilities. Thus we want to allow the same partition of V to have different costs depending on whether it provides information about v_i or $\max_{j\neq i} v_j$.

Strategies and Solution Concept. Buyers have two decisions to make. First, they decide what information to acquire. Then, conditional on their information set, they submit a bid to the seller.

As described above, information acquisition is sequential, and an information strategy consists of two parts. Each buyer i first chooses an information partition about others $\Pi_i^{other} \in \mathcal{P}$. Then, conditional on his information set about others $\pi_i^{other} \in \Pi_i^{other}$,

[15] This is a sufficient statistic for the highest bid faced by i in any symmetric equilibrium, which is what matters for i's gross payoff as that determines his allocation and the price he pays if he wins. That said, it is not without loss, and we investigate the robustness of our results to alternative formulations in Section 6.

he chooses an information partition about his own valuation $\Pi_i^{self} : 2^V \longrightarrow \mathcal{P}$. Finally, each buyer chooses a measurable bidding strategy $\beta_i : 2^V \times 2^V \times U \longrightarrow \mathbb{R}_+$, which outputs a bid given the buyer's overall information $\pi_i = (\pi_i^{other}, \pi_i^{self}, u_i)$.

Let $\Sigma \equiv \left\{(\Pi_i^{other}, \Pi_i^{self}, \beta_i) \in \mathcal{P} \times \mathcal{P}^{2^V} \times \mathbb{R}_+^{2^V+2^V+U}\right\}$ be the overall set of pure strategies,[16] and $\sigma_i \in \Delta\Sigma$ a strategy for buyer i.

Buyer i's ex-ante expected utility under strategy profile $(\sigma_i)_i$ writes

$$\mathbb{E}_{\sigma_i,\sigma_{-i}}\left[U(\nu_i, \beta_i(\pi_i), \beta_{-i}(\pi_{-i})) - \lambda\left(c\left(\Pi_i^{other}, p_{1:N-1}\right) + c\left(\Pi_i^{self}\left(\pi_i^{other}\right), p\right)\right)\right],$$

where $\lambda > 0$ is a parameter that scales the cost of information, and $p_{1:N-1}$ is the prior distribution of $\max_j \tilde{v}_j$.

A Nash equilibrium is a strategy profile such that each buyer's equilibrium strategy σ_i maximizes his ex-ante expected utility given that all others follow theirs. As usual in a second-price auction, there exist many unappealing Nash equilibria, and even more so now that the information structure is endogenous. In particular, the concept of Nash equilibrium imposes no discipline on which (losing) bid a buyer submits if, at some information set, he can predict his toughest opponent's bid and knows that he does not want to outbid it. To rule out unrealistic equilibria, we use the following trembling-hand-like refinement.

Definition 1. *A tremble-robust symmetric equilibrium σ^* is a Nash equilibrium in which,*

(i) all buyers choose the same strategy $\sigma_i = \sigma^*$,

(ii) there exists a sequence of strictly positive numbers $\{\varepsilon^{(k)}\}_{k=1}^{\infty}$ converging to zero such that $\sigma^* = \lim_{k\longrightarrow\infty} \sigma^{(k)}$, where $\sigma^{(k)}$ is a symmetric Nash equilibrium of a perturbed game in which, with probability $\varepsilon^{(k)}$, each buyer's bid is drawn from some continuous distribution F with support $[\min_{\nu_i \in V\times U} \nu_i, \max_{\nu_i \in V\times U} \nu_i]$ independently of his chosen strategy.

Intuitively, we require that the equilibrium remains virtually unchanged if, with vanishing probability, buyers tremble and make a bid drawn from a full-support distribution. This refinement allows us to discipline the bids of losing buyers: if buyer i knows that one of his opponents will bid at least 10 and that he does not want to win at that bid, it ensures that i will bid his expected valuation instead of, e.g., 9.9. This refinement serves no other purpose.

[16] Technically, a buyer could condition his bid not only on what he *learned* π_i but also on the partitions he *chose* $(\Pi_i^{other}, \Pi_i^{self})$—i.e., not only the the signal *realizations* but on the signals themselves. However, that would never be strictly optimal as π_i is a sufficient statistic for a buyer's posterior belief.

Assumptions on the Cost of Information. First, we assume that the cost of a signal only depends on the chosen partition through its effect on the buyer's belief. Any partition $\Pi = \{\pi_l\}_{l=1,\dots L}$ induces a distribution over posterior beliefs, which puts weight on as many posteriors as there are elements in the partition $\{\mu_l\}_{l=1,\dots L}$ with[17]

$$\mu_l(v) = \begin{cases} \frac{\Pr(\tilde{v}=v)}{\sum_{v' \in \pi_l} \Pr(\tilde{v}=v')} \text{ if } v \in \pi_l \\ 0 \text{ otherwise} \end{cases}.$$

We suppose that the cost of information only depends on the chosen partition through the extent to which it reduces the amount of uncertainty in the buyer's belief. Formally, there exists a measure of uncertainty $H : \Delta V \longrightarrow \mathbb{R}_+$, which is a concave function of a belief, such that

$$c(\Pi, \text{prior}) = H(\text{prior}) - \mathbb{E}(H(\text{posterior}) \mid \Pi).$$

We furthermore assume that H is bounded and continuous. This formulation precludes that the cost of a signal about others be greater (or lower) than the cost of a signal about self per se. We could relax this assumption, and say that the cost of a signal about others is scaled by a different parameter λ^{other} than one about self λ^{self}. All our results would go through for $\lambda^{other}/\lambda^{self}$ not too large. More importantly, the cost of leaning about the competition might scale with the number of competitors $N-1$, and we discuss how that would affect our results in Section 6.

A notable example of a cost function that has such a form is the entropic cost.

Example 1. *Let H be the (extended) entropy function $H(p) = -\sum_v p(v)\log[p(v)]$.[18] The cost of an information partition equals the expected reduction in the entropy of the buyer's belief:*

$$c(\Pi, \Pr(\cdot)) =$$
$$-\sum_v \Pr(\tilde{v}=v)\log[\Pr(\tilde{v}=v)] + \sum_{\pi^l \in \Pi} \Pr(\tilde{v} \in \pi^l) \sum_v \frac{\Pr(\tilde{v}=v)}{\Pr(\tilde{v} \in \pi^l)} \log\left(\frac{\Pr(\tilde{v}=v)}{\Pr(\tilde{v} \in \pi^l)}\right).$$

[17] In what follows, the random variable $\tilde{v}$ is either a buyer's own valuation $\tilde{v}_i$ or the valuation of his toughest opponent $\max_{j \neq i} \tilde{v}_j$.

[18] The entropy is usually only defined for full-support beliefs. When a buyer learns $v \in \pi^l \neq V$, he however puts zero probability on all realizations not in π^l: his posterior does not have full support. We extend the domain of the entropy function in the following way. For any belief p on the boundary of the simplex, we define $H(p)$ to be the limit of $-\sum_v \hat{p}(v)\log[\hat{p}(v)]$ as $\hat{p} \longrightarrow p$ for some full-support $\hat{p}$.

The entropic cost has been widely used in the applied literature, in particular in models of rational inattention. It has the advantage of being tractable and having solid information-theoretic foundations.

The key assumption underlying our results is an assumption on the convexity of the cost, or equivalently on the concavity of the measure of uncertainty H.

Assumption 1. *The measure of uncertainty H is strongly concave: there exists $m > 0$ such that, for all $q, q' \in \Delta V$ and all $t \in [0, 1]$,*

$$tH(q) + (1-t)H(q') - H(tq + (1-t)q') \leq -\frac{1}{2}mt(1-t)||q - q'||^2.$$

Since H is concave, the left-hand side is always weakly negative. For H to be *strongly* concave, it must be sufficiently negative: the growth rate of H must have a quadratic upper bound. This assumption ensures that the cost of a partition is sufficiently convex in the fineness of the partition. In particular, it implies that it is cost-efficient for buyers to acquire *some* information about the competitors (i.e., choose an informative, though fairly coarse, partition Π^{other}) to avoid having to become *fully* informed about their own valuations, whenever the prior p is sufficiently uncertain, and m is sufficiently large.

To formalize this, let $\Pi^{other}_{v^*} \equiv \{\{v : v \leq v^*\}, \{v : v > v^*\}\}$ be the partition that divides the set of valuations V into two subsets: valuations that are below some threshold v^* and those that are above. Arguably, this is a fairly coarse partition whenever the set of valuations V is rich. If the equilibrium is efficient, a buyer i who learns $\max_j v_j > v^*$ has little to no incentive to learn to distinguish all valuations $v_i \leq v^*$ since he loses the auction at all of these. Let $\Pi^{self}_{>v^*} \equiv \{\{v : v \leq v^*\}, \{v\}_{v>v^*}\}$ be the partition that bundles all these lower valuations together. This partition is less costly than becoming fully informed of v_i, and potentially significantly so. Similarly, when $\max_j v_j \leq v^*$, buyer i might not want to distinguish all values $v_i > v^*$, and let $\Pi^{self}_{\leq v^*} \equiv \{\{v\}_{v \leq v^*}, \{v : v > v^*\}\}$.[19] Hence even a coarse signal about others can significantly reduce how finely a buyer should learn about his own valuation.

Lemma 1. *There exists $\overline{T}$ and $\overline{m}$ such that if $\sum_v [p(v)]^2 \leq \overline{T}$ and $m \geq \overline{m}$, then it is less costly for buyers to acquire* some *information about others instead of becoming fully informed about*

[19] Such partitions are sometimes called lower and upper censorship in the literature.

themselves. Formally,

$$c\left(\Pi^{other}_{v^*}, p_{1:N-1}\right) + \left(\Pr(v_i \leq v^*)\right)^{N-1} c\left(\Pi^{self}_{\leq v^*}, p\right) + \left[1 - \left(\Pr(v_i \leq v^*)\right)^{N-1}\right] c\left(\Pi^{self}_{>v^*}, p\right)$$
$$< c\left(\left\{\{v_i\}_{v_i \in V}\right\}, p\right),$$

for some $v^* \in V$.

Proofs of all the results are in Appendix B. In Lemma 1, $\sum_v [p(v)]^2$ captures the precision of the prior belief. Indeed, this sum is highest when the prior is deterministic (i.e., when $p(v) = 1$ for some v), and is lowest under a uniform prior. To see why the condition on the prior is necessary, take the extreme case in which only two valuations have strictly positive prior probability: $V = \{\underline{v}, \overline{v}\}$. Then any information acquired about others must fully reveal $\max_j v_j$: the only non-trivial information partition is the fully revealing one $\{\{\underline{v}\}, \{\overline{v}\}\}$. There is then no scope for buyers to save on information costs about their own values by learning a bit about others.

For the rest of the paper, we assume that the prior p is sufficiently uncertain and H is sufficiently concave that Lemma 1 holds.

Example 1 (continued)**.** *Consider again the entropy function $H(p) = -\sum_v p(v)\log[p(v)]$. Let $N = 2$ and $V = \{\frac{1}{K}, \frac{2}{K}, \ldots, \frac{K}{K}\}$ with $p(v_i) = \frac{1}{K}$ for all $v_i \in V$. The fully revealing partition $\Pi^{self} = \{v_i\}_{v_i \in V}$ costs*

$$c\left(\{v_i\}_{v_i \in V}, p\right) = H(p) = \log(K).$$

Suppose K is even. Learning whether the competitor's value v_j is above or below $v^* = \frac{1}{2}$ costs

$$c\left(\Pi^{other}_{v^*}, p_{1:N-1}\right) = \log(K) - \frac{v^*}{K}\log(v^*) - \frac{K - v^*}{K}\log(K - v^*) = \log(2).$$

When buyer i learns v_j is greater (respectively, lower) than $v^* = 0.5$, he only needs to learn his valuation precisely when it is also greater (resp., lower) than $v^* = 0.5$, which costs

$$c\left(\Pi^{self}_{\leq v^*}, p\right) = c\left(\Pi^{self}_{>v^*}, p\right) = \log(K) - \frac{1}{2}[\log(K) - \log(2)].$$

Hence that allows him to reduce learning costs about his value by $\frac{1}{2}[\log(K) - \log(2)]$, and so learning about the competition is cost-efficient whenever $K > 8$.

3 How Competition Shapes Buyers' Information

This section investigates how the competitive pressure between buyers affects what information they seek and the resulting equilibrium information structure. We first consider two benchmark cases in which buyers are either exogenously informed of their valuations or can only acquire costly information about their own valuations. We show that these two benchmarks yield the same predictions when information costs are small relative to the value of the good. This is, however, not the case when buyers can also learn about their competitors.

3.1 Two Benchmark Cases

In the first benchmark we consider, buyers are exogenously informed of their valuations. This case is well understood, and bidding truthfully is a dominant strategy for buyers. Whether or not they know others' valuations, or can acquire information about them, is then irrelevant. We report the properties of the equilibrium for completeness.

Proposition 0. *Suppose buyers know their valuations ex-ante. Then there exists a symmetric equilibrium in which expected revenue equals the expected second-highest valuation* $\mathbb{E}[\nu_{(2)}]$.

Now suppose buyers have no private information ex-ante and can only acquire information on their own valuations. Most papers on information acquisition in auctions focus on this case.

Proposition 1. *Suppose buyers can only learn about themselves. Then, for* λ *small enough, there exists a symmetric equilibrium in which they all become fully informed about their own valuations, and expected revenue equals the expected second-highest valuation* $\mathbb{E}[\nu_{(2)}]$.

Hence, the two benchmarks yield similar predictions for small information costs. The intuition is direct: the gains associated with distinguishing two realizations of $\tilde{v}_i$ are always strictly positive, as a buyer might face a price (i.e., a highest bid) that falls precisely between these two realizations. If information costs are small enough, buyers must choose the fully revealing partition $\Pi^{self} = \{\{v\}_{v \in V}\}$.

3.2 Multi-Dimensional Learning

We now consider our main model specification, in which buyers can acquire information on their valuations as well as others'. We show that buyers have an incentive to learn a bit about their competitors, so as not to waste resources learning about their own valuations when such information makes no difference.

We start by establishing equilibrium existence.

Proposition 2. *There exists a symmetric equilibrium that is robust to trembles for any cost parameter λ.*

In what follows, we use the term "equilibrium" to refer to a tremble-robust symmetric equilibrium.

We now show that, contrary to our benchmarks, buyers cannot all become fully informed of their valuations in equilibrium. This is true even as the cost parameter λ becomes arbitrarily small.

Proposition 3. *Suppose buyers can learn both about themselves and others. Then there does not exist a sequence of equilibria in which buyers become fully informed of their valuations as the cost parameter λ goes to zero. That is, there exists $\varepsilon > 0$ such that, for any sequence of equilibria $\{\sigma_\lambda\}_\lambda$,*

$$\lim_{\lambda \longrightarrow 0} \Pr\left(\Pi^{self} = \{\{v_i\}_{v_i \in V}\} \mid \sigma_\lambda\right) \leq 1 - \varepsilon.$$

The intuition is the following. If buyers learn their valuations fully, they simply bid truthfully in equilibrium, and the good goes to the highest-valuation buyer. It is then cost-efficient for buyers to first assess how much competitive pressure they will face in the auction (i.e., what is the highest valuation among their competitors) and then only learn their own valuations when it is worth, as this leads to strictly lower overall information costs (Lemma 1). In the proof, we show that doing so does not harm their gross payoff from the auction and must hence be a profitable deviation.

More generally, this highlights buyers' incentive to learn about their competitors. If they do so, then their private information (i.e., their types) when entering the auction will be interdependent. Indeed, not only will buyers have information relevant to others, but their beliefs about their own valuations will depend on what they learned about others. In other words, the equilibrium information structure will fail to satisfy the standard assumption of independent private types, despite buyers' valuations being statistically independent. This significantly complicates the analysis of equilibrium behavior, as buyers no longer have a dominant strategy when deciding how to bid.

To illustrate this, consider a buyer i who acquired no information whatsoever and let $N = 2$. If buyers were not able to acquire information about others, then bidding his expected valuation would be a dominant strategy for i. This is no longer the case in our model. For instance, suppose buyer j becomes full informed about his competitor's valuation (i.e., $\Pi_j^{other} = \{\{v\}_{v \in V}\}$) and always bids just below it (i.e., $\beta_j(\pi_j) = v_i + \underline{u}$). Given j's strategy, buyer i *always* wants to win the auction since the price he pays always lies strictly below his valuation: a best response is for i to bid sufficiently high so as to be guaranteed a win. Bidding his expected valuation given *his own information set* results in a strictly lower payoff for i, and so it is no longer a dominant strategy.[20]

Overall, buyers' ability to learn about each other can create rich interdependencies between their information and bids, potentially expanding widely the types of behavior sustainable in equilibrium. Yet, focusing on high-stake auctions (i.e., auctions with small cost parameter λ) enables us to characterize equilibrium information structures.[21]

Theorem 1. *There exists $\overline{\lambda} > 0$ such that, for all $\lambda \leq \overline{\lambda}$, there exists $\varepsilon(\lambda) > 0$ with $\lim_{\lambda \longrightarrow 0} \varepsilon(\lambda) = 0$ such that, if an information structure has probability $\Pr\left(\Pi^{other}, \Pi^{self} \mid \sigma_\lambda\right) \geq \varepsilon(\lambda)$ in some equilibrium σ_λ, then it solves*

$$\min_{\widehat{\Pi}^{other}, \widehat{\Pi}^{self}} \quad c\left(\widehat{\Pi}^{other}, p_{1:N-1}\right) + \mathbb{E}_{\hat{\pi}^{other}}\left[c\left(\widehat{\Pi}^{self}(\hat{\pi}^{other}), p\right)\right]$$

s.t.

($\star$)

$$\widehat{\Pi}^{self}(\hat{\pi}^{other}) = \left\{\left\{v_i \mid v_i < \min_{v \in \hat{\pi}^{other}} v\right\}, \left\{v_i\right\}_{v_i \in \hat{\pi}^{other}}, \left\{v_i \mid v_i > \max_{v \in \hat{\pi}^{other}} v\right\}\right\} \; \forall \hat{\pi}^{other} \in \widehat{\Pi}^{other}.$$

In words, condition ($\star$) requires that if buyer i knows his toughest competitor's valuation belongs to some interval π^{other}, then buyer i learns his own valuation fully if and only if it falls in the same interval. Acquiring more information that this does not affect the outcome of the auction in equilibrium, and thus serves no purpose. Figure 2 illustrate condition ($\star$).

There exist many information structures satisfying ($\star$),[22] but an equilibrium information structure must furthermore minimize total information costs. Intuitively, there is a certain amount of information that guarantees buyers make no mistake at the bid-

[20] Thus, the (more standard) refinement of weakly undominated strategy has little bite in our setting, which is why we require robustness to trembles.

[21] Focusing on high-stake auctions ensures that any information that has non-trivial value must be acquired in equilibrium, thus reducing the noise in buyers' behaviors.

[22] For instance, not acquiring any information about the competition $\Pi^{other} = \{V\}$ and fully learning one's own value $\Pi^{self}(\{V\}) = \{\{v\}_{v \in V}\}$.

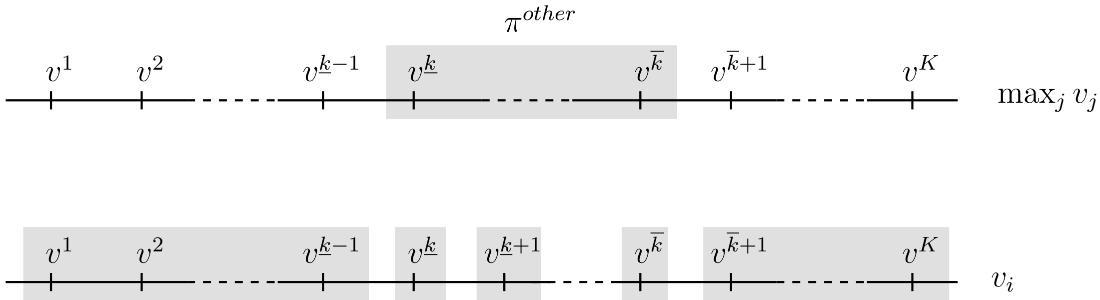


Figure 2: *Let* $V = \{v^1, v^2, \ldots, v^K\}$ *and order the valuations in increasing order, i.e.,* $v^k < v^{k+1}$*. If buyer* i *learns that* $\max_j v_j \in \pi^{other} \equiv \{v : v^{\underline{k}} \leq v \leq v^{\bar{k}}\}$ *(top line), he chooses to fully learn his valuation if it belongs to the set* π^{other}*, but fails to distinguish all valuations that are for sure lower or higher than* $\max_j v_j$ *(bottom line).*

ding stage (condition ($\star$)) and buyers choose the cheapest way to achieve it. Lemma 1 then implies $\Pi^{other} \neq \{V\}$, as fully learning one's own valuation is not cost-efficient, and buyers acquire *some* information about others in equilibrium. Overall, the competitive pressure that buyers impose on each other shapes the information that they acquire in significant ways. The rest of the paper examines how that, in turn, affects the value of competition.

4 Revenue and Entry Distortions

In this section, we analyze the impact of learning incentives on revenue and entry. We show that the equilibrium information structure leads buyers to compete less aggressively for the good, which depresses revenue and distorts entry.

4.1 Revenue Loss

Since buyers do not learn their valuations fully in equilibrium, expected revenue is likely to be different than in our benchmark cases. Our first main result states that revenue remains strictly lower and bounded away from the expected second-highest valuation, even for small information costs λ.

Theorem 2. *Let* $N \geq 3$*. There exist* $L > 0$ *and* $\bar{\lambda} > 0$ *such that, for all* $\lambda \leq \bar{\lambda}$*, the revenue generated in any equilibrium* σ_λ *of the second-price auction is bounded away by* L *from the* *expected second-highest valuation:*

$$\mathbb{E}\left[\textit{ equilibrium revenue }\middle|\sigma_\lambda\right] < \mathbb{E}\left[\nu_{(2)}\right] - L.$$

Note that the constant L is independent of the cost parameter λ. Hence revenue is bounded below the expected second-highest valuation, and does not converge to it as information costs vanish. This contrasts with our above benchmarks, where revenue converges to the expected second-highest valuation as λ goes to zero.

What is driving the loss in revenue? In equilibrium, buyers first assess the competition and only learn their own valuations if they fall in a similar range as that of their toughest competitor. As a result, losing bidders often only learn that their valuations are below some threshold (and, in particular, below that of their toughest competitor) but fail to learn it exactly. Our equilibrium refinement guarantees that if losing bidders fail to learn their valuations, they bid their expected valuations given their information sets. This reduces the variance in losing bids and distorts the expected second-highest bid downwards whenever $N \geq 3$. Indeed, since the $\max$ is a convex function, the expected highest bid among losing bids is greater when losing bids are more dispersed. (With only $N = 2$ buyers, there is only one losing bid, and dispersion plays no role.)

We emphasize that for sufficiently small information costs λ, the equilibrium allocation of the good remains efficient in our model, namely the highest valuation bidder wins the good. Indeed, a buyer only fails to learn his valuation in equilibrium if he is sure of losing (or winning) given what he learned about others. Hence a direct corollary of Theorem 2 is that endogenous information acquisition does not affect total surplus for small information costs λ, but redistributes surplus from the seller to the buyers. It furthermore suggests that, if possible, high-valuation buyers have a strong incentive to signal that they have a high valuation, so as to discourage others from learning about their own and competing aggressively. This is often seen in practice. For instance, jump bidding and toeholds are sometimes seen as signaling devices that aim at deterring competition (Bulow et al. (1999); Betton and Eckbo (2000); Hörner and Sahuguet (2007)).

A Uniform Example. Let $V = \left\{\frac{1}{K}, \frac{2}{K}, \ldots, \frac{K-1}{K}, \frac{K}{K}\right\}$ be the set of possible valuations, and $\Pr(\tilde{v}_i = v) = \frac{1}{K}$ be the prior probability of each $v \in V$. For K large enough, this approximates a uniform distribution on $[0, 1]$. We set H to be the entropy function as in Example 1, such that $H(p) = -\sum_v p(v)\log(p(v))$ for any belief $p \in \Delta V$.

We know from our analysis that for small enough cost parameter λ, buyers only put non-trivial weight on cost-minimizing information structures satisfying ($\star$). We find this information structure numerically,[23] and depict the equilibrium information partition about others Π^{other} for several values of N in Figure 8 of Appendix A.1. We then simulate equilibrium bids and compute expected revenue for vanishing λ.[24] We also compute expected revenue when buyers are fully informed of their valuations, which equals the expected second-highest valuation.

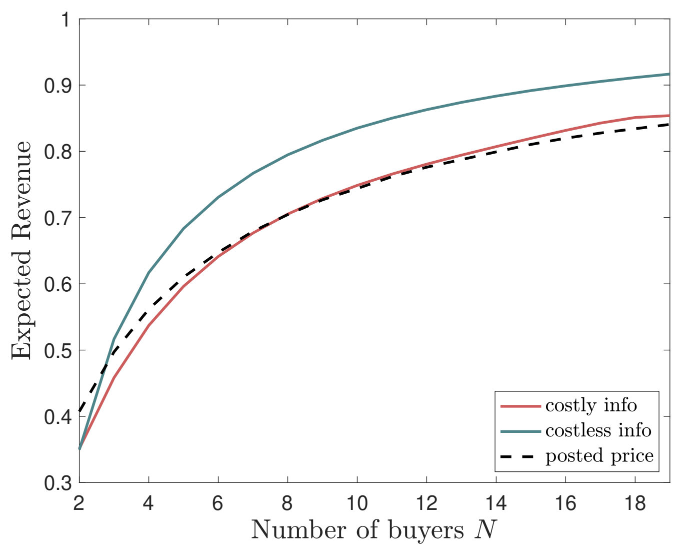
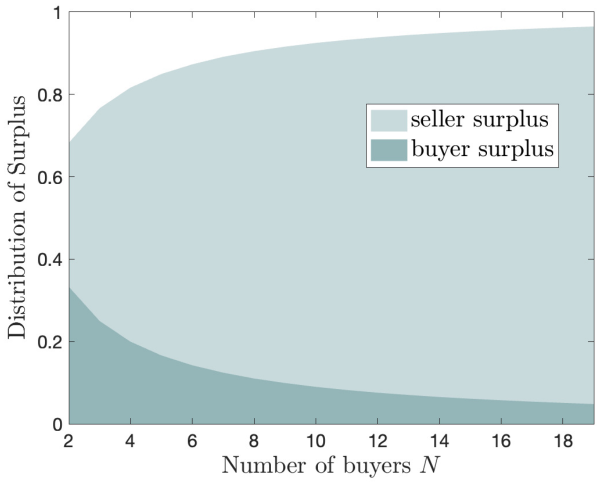


Figure 3: *The top (blue) line plots expected revenue in the standard model (i.e., the expected second-highest value). The bottom (red) line plots expected revenue in our model for small λ. For comparison, the dashed (black) line plots expected revenue in the standard model when the seller uses a posted-price mechanism (i.e., commits to a price and, if more than one buyer is interested, the winner is chosen uniformly at random). The difference between the blue and red lines is the revenue loss from Theorem 1. Parameter $K = 30$.*

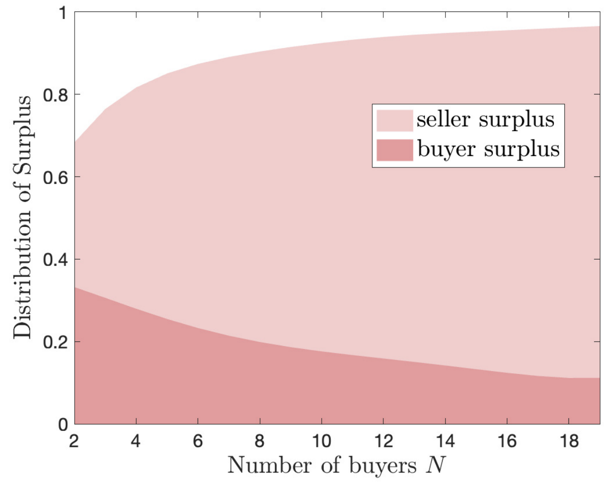
Figure 3 shows how expected revenue remains bounded away from its full information benchmark even as information costs become arbitrarily small. The revenue loss due to endogenous information acquisition is captured by the difference between

[23]Though there could technically exist several cost-minimizing information structures, this never happens in our numerical example.

[24]In Figure 3, we fit a flexible (6^{th}-degree) polynomial on expected revenue to smooth small irregularities arising from the discreteness of the set of valuations V. Indeed, its discreteness induces slight non-monotonicities in N that disappear as the grid of possible valuations gets finer (i.e., as K increases).



(a) Costless Information ($\lambda = 0$)



(b) Costly Information ($\lambda > 0$)

Figure 4: *Distribution of surplus between the seller and buyers in our uniform example. As before, parameter $K = 30$.*

the top (blue) and the bottom (red) lines. To assess the magnitude of this loss, we compare it to the loss in revenue associated with using a posted-price mechanism instead of an auction in the standard model.[25] In this example, the revenue loss due to endogenous information acquisition is similar in magnitude to the loss associated with using a suboptimal posted-price mechanism (difference between blue and dashed black lines).

Since the equilibrium allocation remains efficient when λ is small, the revenue loss means that buyers get a higher surplus than in the standard model with exogenous information. This is illustrated in Figure 4.

4.2 Entry Distortion

We now extend our baseline model to include entry decisions. After the information acquisition stage, buyers decide whether or not to participate in the auction. If they do, they incur an entry cost $\kappa > 0$. Formally, a buyer enters the auction and pays the entry cost κ at information set π_i if he makes a non-zero bid, that is, if $\beta_i(\pi_i) > 0$.

Entry costs are common in practice. For instance, to participate in the 1991 auction for television franchises in the UK, TV channels had to provide a detailed listing of

[25]Under a posted-price mechanism, the seller chooses a (unique) price and, if several buyers express interest in buying the good at that price, allocates the good randomly between them.

what they would air. This resembles more a fixed entry fee than an information acquisition cost. More generally, participating in an auction always entails some fixed (e.g., legal) costs.

Consider first what happens in the standard model where buyers know their valuations ex-ante. In any equilibrium satisfying the tremble-hand-like refinement, buyers who enter the auction bid truthfully. Buyers with higher valuations have a greater incentive to enter, and in equilibrium, a buyer enters if and only if his valuation is above a threshold $\nu_i \geq \nu^*$. A buyer with a valuation precisely at the threshold must be indifferent between entering or not. He knows that he will win the auction only if no one else enters, in which case he pays a price of zero. The indifference condition is then:

$$\Pr\left(\max_{j\neq i} \nu_j < \nu^*\right) \nu^* = \kappa.$$

Let $\tilde{N}_0 = |\{i \mid \tilde{\nu}_i \geq \nu^*\}|$ be the (random) number of buyers who enter the auction in that equilibrium.

If buyers can only learn about themselves, then for λ small enough, there exists an equilibrium in which buyers learn to distinguish all valuations they can have above ν^*, and only enter when their valuations are above ν^*. Hence the equilibrium yields the same allocation and revenue as what the standard model predicts. On the contrary, entry decisions are much different when buyers can also learn about their competitors' valuations.

Proposition 4. *Let $N \geq 3$. There exists $\overline{\lambda} > 0$ such that, for all $\lambda \leq \overline{\lambda}$, there exists $\varepsilon(\lambda) > 0$ with $\lim_{\lambda \longrightarrow 0} \varepsilon(\lambda) = 0$ such that, in any equilibrium σ_λ, the probability that at least two buyers enter the auction is bounded above by*

$$\Pr\left(\tilde{N}_\lambda > 1 \mid \sigma_\lambda\right) \leq \Pr\left(v_{(1)} = v_{(2)}\right) + \varepsilon(\lambda),$$

where $\tilde{N}_\lambda = |\{i \mid \beta_\lambda(\tilde{\pi}_i) > 0\}|$ is the (random) number of buyers who enter.

Proposition 4 states that several buyers enter the auction only if their valuations fall in a similar range. In the proof, we show that two buyers i, j, with values $v_i < v_j$ cannot *both* enter the auction with non-vanishing probability in equilibrium. If they were to both enter, buyer i would never win the auction as his overall valuation ν_i must lie strictly below j's. He would then want to learn j's valuation so as not to enter in those states and save on the entry cost. Entry costs then reinforce the revenue loss described in Theorem 2, as losing buyers not only fail to learn their valuations for the good but stay out of the auction altogether. This is precisely what happened in the 1991 U.K. auction for television franchises in two regions (the Midlands and Scotland). In each region, the incumbent firm was expected to win and ended up being the only one putting together a complete programming plan and bidding for the license.[26]

A direct corollary of Proposition 4 is that, for sufficiently small entry cost κ, there is more entry in the standard model than in ours. Indeed, for κ small enough, the entry threshold when buyers know their valuations ν^* is very close to the lowest possible valuation. With probability close to one, all buyers then enter the auction. On the contrary, the probability that two or more buyers enter the auction remains bounded below one in our model (Proposition 4), irrespective of the size of the entry cost κ.

Observe that, once entry decisions are introduced, the standard model (in which buyers know their own valuations ex-ante, but not that of others) yields different predictions from a model with costless information ($\lambda = 0$, such that buyers know both their own and their toughest competitor's valuations). Indeed, in the former, entry decisions are characterized by the above threshold ν^*. In the latter, two buyers with valuations $v_i \neq v_j$ cannot both enter. Unlike for revenue, there is then no discontinuity in entry decisions between a model in which information is costly ($\lambda > 0$ but small) and one in which it is free ($\lambda = 0$).

5 Market Design Implications

So far, we have shown that losing buyers often fail to learn their valuations precisely, which leads them to bid less aggressively and depresses expected revenue. This has implications for market design. First, the value of an optimal reserve price often dominates the value of an additional bidder. This contrasts with common wisdom inherited from Bulow and Klemperer (1996) (thereafter, "BK"). Second, the seller gains by maintaining uncertainty over competition. Overall, this suggests that competition is most effective if it is carefully *designed* by the seller. We emphasize that the following results provide practical insights on auction design, but rely on stricter assumptions on the environment beyond what is needed for our main results on equilibrium characterization and revenue loss.

[26] See Klemperer (2002) for more details.

5.1 Additional Buyer vs. Reserve Price

In a seminal paper, BK show that the value of an additional bidder in an auction always dominates the value of optimizing the reserve price. This is an important result as it gives market designers a very simple and actionable insight: attract as many bidders as possible.

We revisit this result in our setting with information acquisition. To do so, we let the support of the noise terms $[\underline{u}_N, \overline{u}_N]$ depend on the number of buyers N.[27] This will be important for the following result and we discuss why afterwards.

Proposition 5. *Suppose $(\overline{u}_N - \underline{u}_N) = o(\exp(-N))$ as $N \longrightarrow \infty$. There exists $\overline{N}$ such that, for all $N \geq \overline{N}$ and for λ small enough, revenue with $N+1$ bidders in the second-price auction without reserve is* ***lower*** *than revenue with N bidders in the second-price auction with optimal reserve.*

That are several forces underlying Proposition 5. First, competition is less valuable than in the standard model as additional buyers fail to learn their valuations fully, and so do not add as much competitive pressure. Second, a reserve price is more valuable in our setting. Indeed, losing buyers often fail to learn their valuations precisely, which leaves a larger expected gap between the highest and second-highest bid, and hence more room for a carefully-designed reserve price to intervene. Of course, the standard force—i.e., that higher N means higher second-order statistic on average—is still present. The crux of the proof lies in showing that the first two dominate for N large enough.

It is important to note that the value of both a reserve price and an additional bidder goes to zero as N goes to infinity. In the proof, we show that the latter goes to zero at a faster rate, such that a reserve price dominates whenever the number of buyers N is sufficiently large. This result is not driven by higher N having a negative impact on revenue—our examples suggest revenue is still increasing in N,—but instead shows that auction design has a *relatively* larger value when information is endogenous.

Why is the condition on the support of the noise terms needed? Think of what happens when the number of buyers gets large. With high probability, at least two buyers will have the highest valuation possible $\max_{v \in V} v \equiv v^K$. All these buyers will

[27] We emphasize that this does not change any of our preexisting results. So far, we have taken N as given, so the support of the noise terms could have very well be indexed by N as well. To save on notation, we however decided to only introduce this indexing now.

learn their valuations fully and bid $\nu_i = v^K + u_i$. In such cases, revenue is effectively the same as in the standard auction model with $|\{i : v_i = v^K\}|$ bidders and valuations drawn from $[v^K + \underline{u}, v^K + \overline{u}]$. In that world, the BK result holds as an additional bidder increases the expected highest draw of u_i. This force goes against our result but does not overturn it whenever the support of the u_is shrinks fast enough with N.[28] The driving force behind our result occurs when at most one buyer has the highest possible valuation $v_i = v^K$. In that case, losing buyers fail to learn their valuations and their (low) bids set the price of the good. A reserve price then induces a significant increase in revenue.

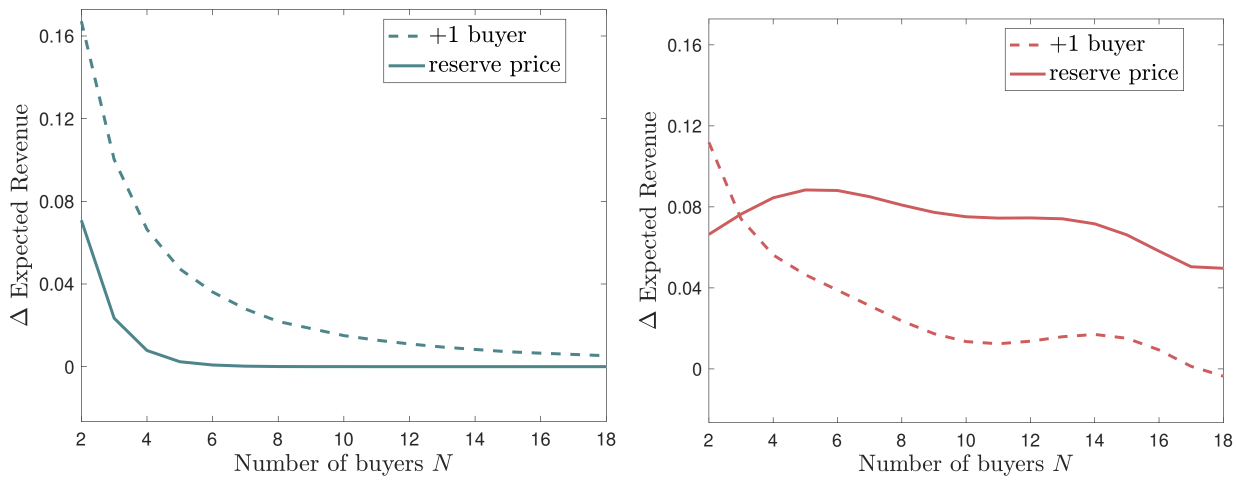
The proof of Proposition 5 requires the number of buyers N to be large enough. We see this as a limitation on the robustness of the result in general. Our uniform example however suggests that N does not have to be very large for Proposition 5 to apply.

Uniform Example Continued. We revisit the uniform example from Section 3.3 to illustrate Proposition 5. The left panel of Figure 5 plots the value of an additional bidder and the value of a reserve price in the standard model, in which buyers know their valuations and bid truthfully. The value of an additional bidder is then always greater than that of a reserve price (BK's result). The opposite is true in our model (right panel).

Not only is the optimal reserve price more valuable in our framework than in standard theory, but it also has different properties. When buyers know their valuations ex-ante, the optimal reserve price is known to be independent of the number of participants N in the auction (see, e.g., Myerson (1981)).[29] In our framework, the optimal reserve price converges to the highest possible valuation as the number of buyers N grows large and, more generally, seems to be increasing in N. See Figure 9 in Appendix A.1 for an illustration.

[28] This assumption does not imply that revenue in the standard model (i.e., $\mathbb{E}[\nu_{(2:N)}]$) is decreasing in N. Indeed, an additional buyer increases the expected second-highest main component $\mathbb{E}[v_{(2:N)}]$ by at least $p(v^K)(1-p(v^K))^N[v^K - \mathbb{E}(v_i \mid v_i < v^K)]$ while the shrinking support of the noise terms induces a loss of at most $\overline{u}_N$.

[29] This result technically requires buyers' set of possible valuations to be an interval and the distribution to satisfy an appropriate regulatory condition. It holds approximately when our finite set of valuations is sufficiently fine.



(a) Costless Information ($\lambda = 0$) (b) Costly Information ($\lambda > 0$)

Figure 5: *The dashed lines plot the value of an additional bidder—i.e., the difference between expected revenue under $N+1$ bidders and expected revenue under N bidders—as a function of N. The solid lines plot the value of a reserve price—i.e., the difference between expected revenue under N bidders with an optimal reserve price and without a reserve price—as a function of N. All curves are fitted with a flexible (6^{th}-degree) polynomial to smooth small irregularities arising from the discreteness of the set of valuations V.*

5.2 Maintaining Uncertainty over Competition

To mitigate the revenue loss, the seller needs to incentivize buyers to learn their valuations for the good. In this section, we show that by inducing uncertainty on the set of buyers allowed into the auction, the seller can induce higher information acquisition and increase revenue.

The set of potential buyers is still exogenous and equal to $N = \{1, \ldots N\}$, but the seller can now commit to only letting some (random) subset of buyers compete in the auction. That is, the seller can commit to only considering some of the submitted bids. Let $\tilde{M}$ be the random set of buyers who get access to the auction—that is, buyers whose bids are taken into consideration,—whose distribution is chosen by the seller. As before, buyers acquire information before bidding and, in particular, before knowing the realization of $\tilde{M}$.

Consider the following way to randomize access to the auction. With probability $1 - q < 1$, all buyers get access $M = \{1, \ldots, N\}$. All bids are then taken into consideration: the good goes to the highest bidder who pays the second-highest bid. With probability q, one buyer chosen uniformly at random is denied access to the auction: $M = \{1, \ldots, N\} \setminus i$ for some i. In such an event, the seller acts as if buyer i had not submitted a bid. Hence, even if one learns that another buyer has a greater valuation, there is still some strictly positive probability q/N that the other buyer's bid will not be accounted for. Information about one's own valuation is then strictly beneficial: for λ sufficiently small, buyers become fully informed.

Proposition 6. *Take any $\varepsilon > 0$. There exists an access rule $\tilde{M}$ such that, for λ small enough,*

$$\mathbb{E}\left[\text{ equilibrium revenue } \middle| \sigma_\lambda\right] \geq \mathbb{E}\left[\nu_{(2)}\right] - \varepsilon$$

in any equilibrium σ_λ. In particular, the access rule described above with $q = 1 - \frac{\varepsilon}{\mathbb{E}[\nu_{(2)}]}$ yields such revenue.

Proposition 6 suggests that randomizing access to the auction is a powerful tool to incentivize information acquisition. By maintaining uncertainty over the competition that a buyer will face in the auction, the seller reduces the negative effect of competition on learning incentives. In the proof of Proposition 6, we show that if a buyer's toughest competitor has a non-zero chance of being excluded from the auction, then the buyer has a strict incentive to learn his valuation for the good. For λ sufficiently small, he will then do so.

Even though this intuition relies on buyers only being able to learn about their toughest competitor $\max_j v_j$—instead of *all* their competitors $(v_j)_j$,—Proposition 6 does not. Indeed, even if buyers could learn about the entire vector $(v_j)_j$, any access rule that excludes all *subsets* of buyers with positive probability—instead of one buyer at a time—would give buyers a strict incentive to learn their own valuations.

Our results can explain why sellers sometimes try to keep secret the identity of participants in an auction. For instance, potential bidders in takeover auctions sign a confidentiality agreement that prevents them from revealing, among other things, their participation in the auction and the value of their bids (see Gentry and Stroup (2019) for a description of a typical takeover auction).

Finally, we compare the value of randomizing access to that of setting an optimal reserve price—which is sometimes a complicated endeavor when the seller has little information about the value of the object.

Proposition 7. *Suppose* $(\overline{u}_N - \underline{u}_N) = o(\exp(-N))$ *as* $N \longrightarrow \infty$*. There exists* $\overline{N}$ *such that, for all* $N \geq \overline{N}$ *and for* λ *small enough, randomizing access to the auction leads to* ***higher*** *revenue than setting an optimal reserve price.*

Hence, when information is endogenous and shaped by competition, randomizing access—and thus sometimes allocating the good to the *wrong* buyer—improves revenue more than setting an optimal reserve price. These two allocative distortions serve different purposes: the former incentivizes buyers to acquire more information about their valuations for the good, while the latter reduces the rent they get from said information. Proposition 7 then says that in our framework, there is more value in incentivizing buyers to acquire information than in reducing their information rent.

6 Discussion and Robustness of the Model

Key assumptions in our analysis are those imposed on the process of information acquisition. Some are required for tractability, while others can be relaxed to some extent. We now discuss these assumptions in more detail.

6.1 The Structure of the Learning Process

We model information acquisition as a two-step process, in which buyers first acquire a signal about their competitors' valuations and then one about their own. The following proposition says that such *ordering* is without loss within the class of two-step learning processes.

Proposition 8 (order of signals)**.** *Consider our main model specification, but suppose that buyers can now choose in which* order *to acquire the two signals.*[30] *There exists* $\overline{\lambda} > 0$ *such that, for all* $\lambda \leq \overline{\lambda}$ *and in any equilibrium* σ_λ*, buyers choose to acquire information on others first, and all results are unchanged.*

The proof of Proposition 8 consists in showing that, in any equilibrium, buyers choose to acquire information about the competition first. Hence imposing that particular ordering is without loss for our results.

[30] That is, they can choose whether to first acquire a signal about their own value v_i and then one about others $\max_j v_j$, or vice versa.

Only looking at two-step learning processes is a stronger assumption. In general, there is no reason why buyers would not want to undertake a more elaborate learning process in which they go back and forth between acquiring some information about their own values and some information about others. We show that giving buyers this added flexibility does not change the main results of the paper: buyers learn about the competition and equilibrium revenue remains bounded below the expected second-highest valuation as information costs vanish.

Proposition 9 (number of signals)**.** *Consider our main model specification, but suppose that buyers can now choose as many signals—i.e., partitions—as they wish as well as in which order to acquire them.*[31] *Proposition 3 and Theorem 2 hold. That is, as information costs vanish, buyers fail to learn their valuations fully which results in a revenue loss.*

If anything, this added flexibility allows buyers to even better condition how much they learn about themselves on the competition they face. To illustrate this, consider our uniform example with $|V| = 20$ possible valuations, $N = 2$ buyers, and entropic cost. In our baseline model, the equilibrium information structure is as depicted in

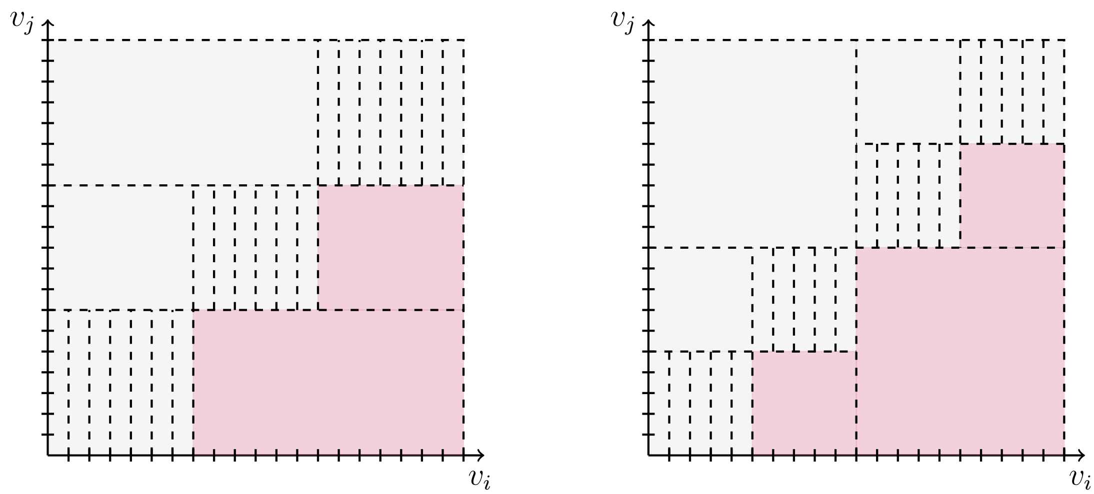


Figure 6: Equilibrium information structure when buyers can only acquire two signals (left panel) vs. as many signals as they wish (right panel).

Figure 6 (left panel): buyers first choose $\Pi^{other} = \{\{v_i : v_i \in [0, \frac{1}{3}]\}, \{v_i : v_i \in [\frac{1}{3}, \frac{2}{3}]\}, \{v_i : v_i \in [\frac{2}{3}, 1]\}\}$ and then learn their valuation fully only if it falls into the same interval

[31]For instance, they can choose to first learn whether $\max_j v_j > v^*$ and then learn whether their own value is above v^*, and then go back and learn more about the competition, etc.

as the opponent's. When buyers are allowed to go back and forth between learning about self and others, the equilibrium information structure is as depicted on the right panel and resembles a binary search: buyers first learn whether $v_j > 0.5$ and whether $v_i > 0.5$, which devides the state space into four quadrants. If $v_j > 0.5$ and $v_i \leq 0.5$ (upper left quadrant) or if $v_j \leq 0.5$ and $v_i > 0.5$ (lower right quadrant), then they stop acquiring information. Otherwise they keep on learning. The areas shaded in pink are all the states of the world in which buyer i is the losing buyer and fails to learn his valuation fully, which drive the revenue loss. If anything, the shaded area is larger when buyers have more flexibility in how they learn.

6.2 Noisy Signals and Ex-Ante Private Information

Throughout the paper, we have modeled signals as information partitions. Partitions are a particular type of signals in that they are "deterministic:" the signal can be partially informative but it can never be "wrong." Naturally, this is a simplifying assumption, but one that is needed for tractability. With more noise in the learning process, characterizing equilibrium bidding becomes much more challenging. Recall that in our setting, the second-price auction is no longer strategy-proof: that is, it might not be optimal for a buyer i to bid his expected valuation given his information set. Indeed, other buyers might have acquired information about their competitors—which include i himself—and so their bids might carry information relevant to i. There is no nice structure on how a buyer should update about his value upon tying at a particular bid—e.g., he could update positively at a some bid but negatively at another. By reducing the noise in the learning process we recover tractability: if a buyer ties, he must have a valuation close to that of his toughest opponent, and so must have learned his valuation fully. This is also why we focus on small information cost λ.

In this section, we tweak our baseline model to allow for more general signal structures. Instead of being completely uninformed of their values at the beginning of the game, buyers are now endowed with *some* private information. More precisely, they get to observe for free the realization of a signal s_i where $s_i = \underline{s}$ when $v_i \leq v^*$ and $s_i = \overline{s}$ when $v_i > v^*$ for some $v^* \in V$. In words, they initially know whether their value is "high" (above v^*) or "low" (below v^*.) As in our baseline model, buyers can acquire two additional signals: one about their own value and one about their competitors'. However, we now impose that buyers can only learn about their competitors' *exoge-*

nous information, i.e, they can only learn about $\max_j s_j$ and not about $\max_j v_j$. This assumption makes all the above complications go away as it ensures that a buyer i can only learn things about j that j *already knows*. We can thus allow for more general signals and we give buyers full flexibility in how to design these two signals. That is, they can choose any distribution over posterior beliefs about others ($\max_j s_j$) and about self (v_i) that is consistent with their prior. The rest of the model is unchanged.

Theorem 3. *If the prior conditional on s_i is sufficiently uncertain,*[32] *then Proposition 3 and Theorem 2 holds.*

Our main results thus extend to a setting where buyers can choose noisy signals: as information costs vanish, buyers still fail to learn their valuations fully, which results in a revenue loss. The same intuition as in our baseline model holds: for small enough information costs, buyers must learn enough so as not to make non-trivial mistakes at the bidding stage. One way not to make mistakes is to learn one's valuation fully, but that is not cost-efficient. In equilibrium, a buyer who knows his value is above v^* (i.e., a buyer with $s_i = \overline{s}$) will converge to learning almost perfectly whether his toughest opponent is weak ($\max s_j = \underline{s}$), in which case he will not acquire any information about himself. Doing so is cost-efficient whenever the remaining uncertainty that buyer i has about his value (i.e., the uncertainty in the distribution of v_i conditional on s_i) is sufficiently large.

6.3 Learning about the max

A perhaps more economically substantive assumption concerns buyers' ability to learn about their toughest competitor's valuation $\max_j v_j$ without having to learn about each competitor $(v_j)_j$ individually. This assumption is mainly made for tractability, as it significantly reduces the dimensionality of the problem. Combined with the assumption on the cost of information, it however does imply that the cost of learning about the competition does not scale with the number of competitors. This seems a good approximation in settings where the number of auction participants in small or where buyers learn about the competition in aggregate—e.g., by undertaking market research. In other settings, however, buyers might only be able to learn about their competitors one at a time. That is, buyers might only be able to learn about $\max_j v_j$ by learning about each v_j independently.

[32] That is, if $\sum_{v \leq v^*}[p(v)/(\sum_{v' \leq v^*} p(v'))]^2$ and $\sum_{v > v^*}[p(v)/(\sum_{v' > v^*} p(v'))]^2$ are sufficiently small.

We now consider this possibility, and investigate how it affects our results in the context of our running uniform example. Suppose that a buyer i wants to learn whether $\max_j v_j > v^*$ for some v^*. What is the cost of getting that information in this new setting? If buyer i can only learn about his rivals *simultaneously*, then he must learn whether each $v_j > v^*$ for all $j \neq i$. The cost of getting that information is then $(N-1)c(\{\{v_j \mid v_j \leq v^*\}, \{v_j \mid v_j > v^*\}\}, p)$ since he must acquire that partition $N-1$ times, once for each opponent. Now suppose that buyer i has a bit more flexibility and can learn about his opponents *sequentially*. So buyer i can learn whether a particular opponent l has $v_l > v^*$, and if yes, he does not need to learn about the others. Thus learning sequentially reduces overall expected learning costs: the cost of learning whether $\max_j v_j > v^*$ is now $\left[\sum_{n=1}^{N-1} \Pr(v_j \leq v^*)^{n-1}\right] c(\{\{v_j \mid v_j \leq v^*\}, \{v_j \mid v_j > v^*\}\}, p)$.

We revisit our uniform example one last time and derive equilibrium revenue under these two possible learning specifications. The results are depicted in Figure 7. Focus first on the top graphs. When buyers must learn about all their opponents at the same time, then it remains cost-efficient to acquire some information about the competition only if the number of competitors is not too large, in which case the revenue loss persists. As N increases, inquiring about the competition becomes increasingly costly, and above some threshold ($N = 8$, here) it is no longer cost-efficient to do so. If buyers can learn about their opponents sequentially, then doing so remains cost-efficient even for relatively large N. Hence, under such alternative modeling, Theorem 1 and 2 persist for auctions of intermediate size but not necessarily for large auctions, depending on whether buyers can learn sequentially.

Proposition 5 however seems less robust to alternative assumptions on the cost of learning. Recall that Theorem 5 compares the value of a reserve price to the value of an additional bidder. Now that the cost of learning about others scales with the number of competitors, having an additional bidder in the auction might discourage bidders from learning about the competition. This force goes against our result and was absent in our baseline model.

7 Conclusion

This paper develops a tractable model of multidimensional information acquisition in auctions, in which buyers can learn both about the strength of competition and about their own valuations. We first characterize how the competitive pressure be-

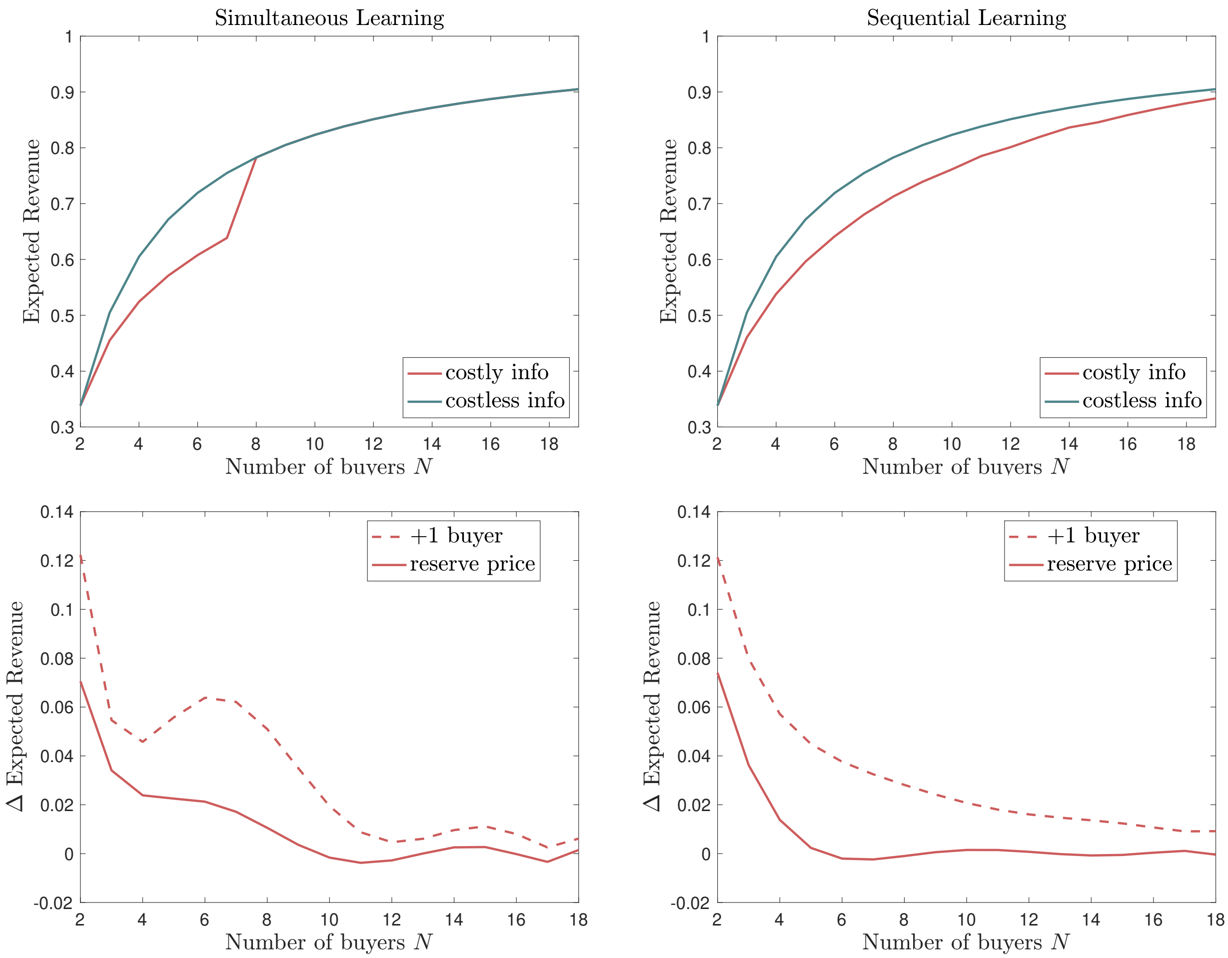


Figure 7: *The top figures plot expected revenue in the standard model (blue) and in our model for small λ (red). The bottom figures plot the value of an additional bidder (dashed line) and the value of a reserve price (solid line) in our model. Both are shown for two different ways buyers can learn about the competition and their associated costs. The left figures consider simultaneous learning while the right figures allow for sequential learning. All curves are fitted with a flexible (6^{th}-degree) polynomial to smooth small irregularities arising from the discreteness of the set of valuations V. Parameter $K = 100$.*

tween buyers shapes the information that they seek. In our framework, buyers find it cost-efficient to first acquire some information about their competitors so as to only learn their valuations when they have a chance to win. Second, we show that competition between buyers is made less effective by learning incentives. Losing buyers often fail to learn their valuations precisely and, as a result, compete less aggressively for the good. This depresses expected revenue. We then propose market design solutions to mitigate these effects. Overall, we show that the seller benefits from carefully designing the competition that buyers face—either via a reserve price or by maintaining uncertainty over the set of auction participants.

Our results suggest that the interactions between information and competition can have large, previously unexplored implications for auction design. We believe it provides yet another justification for robust mechanism design, or at least a careful consideration of informational incentives in the practice of market design.

REFERENCES

Bergemann, D. and Pesendorfer, M. (2007). Information structures in optimal auctions. *Journal of economic theory*, 137(1):580–609.

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

Betton, S. and Eckbo, B. E. (2000). Toeholds, bid jumps, and expected payoffs in takeovers. *The Review of Financial Studies*, 13(4):841–882.

Bobkova, N. (2019). Information Choice in Auctions. Working paper.

Bulow, J., Huang, M., and Klemperer, P. (1999). Toeholds and takeovers. *Journal of Political Economy*, 107(3):427–454.

Bulow, J. and Klemperer, P. (1996). Auctions versus Negotiations. *American Economic Review*, 86(1).

Bulow, J. and Klemperer, P. (2009). Why do sellers (usually) prefer auctions? *American Economic Review*, 99(4):1544–75.

Cantillon, E. (2008). The effect of bidders' asymmetries on expected revenue in auctions. *Games and Economic Behavior*, 62(1):1–25.

Cole, R. A., Ferris, K. R., and Melnik, A. (2016). The direct cost of advice in m&a transactions in the financial sector. *Available at SSRN 1458465*.

Compte, O. and Jehiel, P. (2002). On the value of competition in procurement auctions. *Econometrica*, 70(1):343–355.

Compte, O. and Jehiel, P. (2007). Auctions and information acquisition: sealed bid or dynamic formats? *The Rand Journal of Economics*, 38(2):355–372.

Gentry, M. and Stroup, C. (2019). Entry and competition in takeover auctions. *Journal of Financial Economics*, 132(2):298–324.

Gershkov, A., Moldovanu, B., Strack, P., and Zhang, M. (2021). A theory of auctions with endogenous valuations. *Journal of Political Economy*, 129(4):1011–1051.

Gleyze, S. and Pernoud, A. (2022). Informationally simple incentives. *Working paper*.

Harsanyi, J. C. (1973). Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International journal of game theory*, 2(1):1–23.

Hausch, D. B. and Li, L. (1993). A common value auction model with endogenous entry and information acquisition. *Economic Theory*, 3(2):315–334.

Hörner, J. and Sahuguet, N. (2007). Costly signalling in auctions. *The Review of Economic Studies*, 74(1):173–206.

Jehiel, P. and Lamy, L. (2015). On discrimination in auctions with endogenous entry. *American Economic Review*, 105(8):2595–2643.

Kim, J. and Che, Y.-K. (2004). Asymmetric information about rivals' types in standard auctions. *Games and Economic Behavior*, 46(2):383–397.

Kim, J. and Koh, Y. (2020). Learning rivals' information in interdependent value auctions. *Journal of Economic Theory*, 187:105029.

Kim, K. and Koh, Y. (2022). Auctions with flexible information acquisition. *Games and Economic Behavior*, 133:256–281.

Kleinberg, R., Waggoner, B., and Weyl, E. G. (2016). Descending price optimally coordinates search. *arXiv preprint arXiv:1603.07682*.

Klemperer, P. (2002). What really matters in auction design. *Journal of economic perspectives*, 16(1):169–189.

Larson, K. and Sandholm, T. (2001a). Computationally limited agents in auctions. In *AGENTS-01 Workshop of Agents for B2B*, volume 28.

Larson, K. and Sandholm, T. (2001b). Costly Valuation Computation in Auctions. *Theoretical Aspects of Rationality and Knowledge (TARK VIII)*, pages 169–182.

Levin, D. and Smith, J. L. (1994). Equilibrium in auctions with entry. *The American Economic Review*, pages 585–599.

Lu, J., Ye, L., and Feng, X. (2021). Orchestrating information acquisition. *American Economic Journal: Microeconomics*, 13(4):420–65.

Marquez, R. (2021). Selling assets: Are sellers better off with strong buyers? *Available at SSRN 3949699*.

Maskin, E. and Riley, J. (2000). Asymmetric auctions. *The review of economic studies*, 67(3):413–438.

Miettinen, P. (2013). Information acquisition during a dutch auction. *Journal of Economic Theory*, 148(3):1213–1225.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of operations research*, 6(1):58–73.

Persico, N. (2000). Information Acquisition in Auctions. *Econometrica*, 68(1):135–148.

Prummer, A. and Nava, F. (2023). Value design in optimal mechanisms. Technical report, Working Paper.

Quint, D. and Hendricks, K. (2018). A Theory of Indicative Bidding. *American Economic Journal: Microeconomics*, 10(2):118–51.

Reny, P. J. (1999). On the existence of pure and mixed strategy nash equilibria in discontinuous games. *Econometrica*, 67(5):1029–1056.

Roberts, J. W. and Sweeting, A. (2013). When should sellers use auctions? *American Economic Review*, 103(5):1830–61.

Salant, D. J. (1997). Up in the air: Gte's experience in the mta auction for personal communication services licenses. *Journal of Economics & Management Strategy*, 6(3):549–572.

Shi, X. (2012). Optimal Auctions with Information Acquisition. *Games and Economic Behavior*, 74(2):666–686.

Tian, G. and Xiao, M. (2007). Endogenous information acquisition on opponents' valuations in multidimensional first price auctions.

Ye, L. (2007). Indicative Bidding and a Theory of Two-Stage Auctions. *Games and Economic Behavior*, 58(1):181–207.

APPENDIX A ADDITIONAL MATERIAL

A.1 Additional Details on the Uniform Example

Throughout the paper, we use a uniform example in which $V = \left\{\frac{1}{K}, \frac{2}{K}, \ldots, \frac{K-1}{K}, \frac{K}{K}\right\}$ and $\Pr(\tilde{v}_i = v) = \frac{1}{K}$ for all $v \in V$ to illustrate our results. We know from Theorem 1 that, in equilibrium, all buyers choose the cost-minimizing information structure $(\Pi^{other}, \Pi^{self})$ that satisfies

$$\Pi^{self}\left(\pi^{other}\right) = \left\{ \left\{ v_i : v_i < \min_{v \in \pi^{other}} v \right\}, \left\{ v_i \right\}_{v_i \in \pi^{other}}, \left\{ v_i : v_i > \max_{v \in \pi^{other}} v \right\} \right\},$$

for all $\pi^{other} \in \Pi^{other}$. We find this cost-minimizing information structure numerically, and depict Π^{other} for several values of N in Figure 8. For instance, when $N = 3$, Π^{other} partitions the set of valuations into four intervals: buyers learn whether their toughest competitor has a valuation below $.28$, between $.28$ and $.5$, between $.5$ and $.72$, or above $.72$.

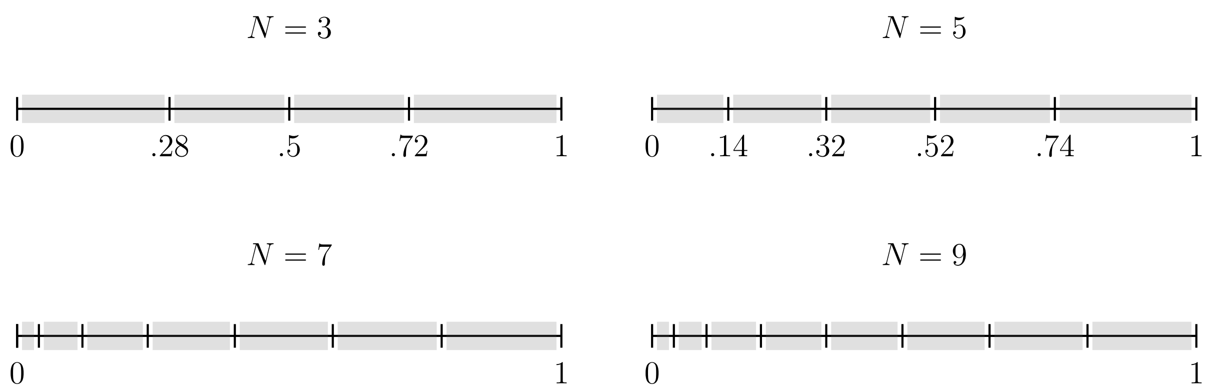


Figure 8: *Parameter* $K = 50$.

We furthermore compute the optimal reserve price, both in our framework and when buyers are exogenously informed of their valuations, and plot them in Figure 9. In our framework, the optimal reserve price seems increasing in the number of buyers N and we can show that it converges to the highest possible valuation as N goes to infinity. In the standard model, the optimal reserve price is approximately constant (the slight variations are solely driven by the finiteness of the set of valuations).

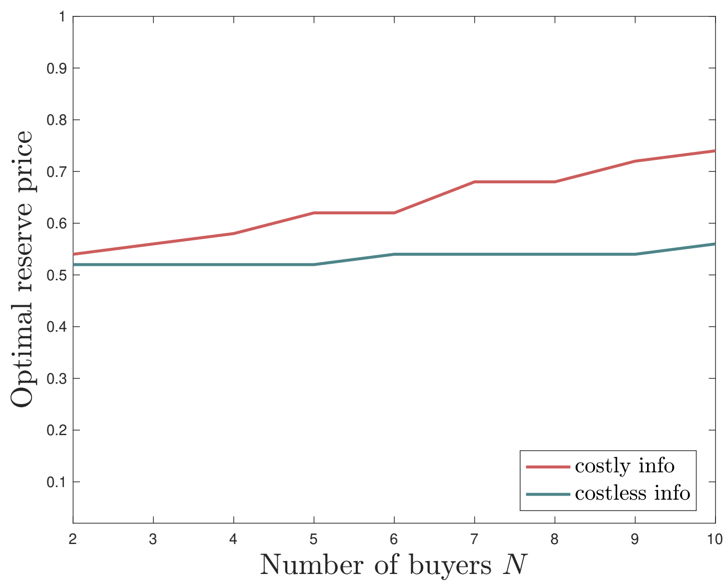


Figure 9: *The optimal reserve price under exogenous information (blue line) is close to being independent of* N*. The fact that it varies slightly with* N *comes from the discreteness of* V*, and disappears for* K *large enough. On the contrary, the optimal reserve price when information is endogenous and shaped by competition is increasing in* N*. Parameter* $K = 50$*.*

Appendix B Proofs

B.1 Preliminary Analysis

Proof of Lemma 1. We want to show that, for $\sum_v [p(v)]^2$ low enough and m large enough, there exists $v^* \in V$ such that

$$c\left(\Pi^{other}_{v^*}, p_{1:N-1}\right) + \left(\Pr(v_i \le v^*)\right)^{N-1} c\left(\Pi^{self}_{\le v^*}, p\right) + \left[1 - \left(\Pr(v_i \le v^*)\right)^{N-1}\right] c\left(\Pi^{self}_{>v^*}, p\right) < c\left(\left\{\{v_i\}_{v_i \in V}\right\}, p\right).$$

Acquiring partition $\Pi^{other}_{v^*}$ leads agent i to hold one of two possible posterior beliefs about $\max_j v_j$. Let $\delta_{v_{1:N-1} \le v^*}$ denote i's posterior about $\max_j v_j$ after learning $\max_j v_j \le v^*$, that is

$$\delta_{v_{1:N-1} \le v^*} \equiv \Pr(\cdot \mid \max_j v_j \le v^*) = \mathbb{1}\{\cdot \le v^*\} \frac{p_{1:N-1}(\cdot)}{\Pr(\max_j v_j \le v^*)}.$$

Similarly, let $\delta_{v_{1:N-1} > v^*}$ denote i's posterior about $\max_j v_j$ after learning $\max_j v_j > v^*$.

Using the fact that the cost of a partition equals the expected reduction in uncertainty in a buyer's beliefs, the above inequality rewrites as

$$\begin{aligned} & H(p_{1:N-1}) - \Pr(\max_j v_j \le v^*) H\left(\delta_{v_{1:N-1} \le v^*}\right) - \Pr(\max_j v_j > v^*) H\left(\delta_{v_{1:N-1} > v^*}\right) \\ & + \Pr(\max_j v_j \le v^*) \left[H(p) - \Pr\left(v_i > v^*\right) H\left(p\left(\cdot \mid v_i > v^*\right)\right) - \sum_{v \le v^*} p(v) H(\delta_v) \right] \\ & + \Pr(\max_j v_j > v^*) \left[H(p) - \Pr\left(v_i \le v^*\right) H\left(p\left(\cdot \mid v_i \le v^*\right)\right) - \sum_{v > v^*} p(v) H(\delta_v) \right] \\ & \qquad < H(p) - \sum_v p(v) H(\delta_v). \end{aligned}$$

Simplifying terms, this becomes

(1)

$$H(p_{1:N-1}) - \Pr(\max_j v_j \le v^*) H\left(\delta_{v_{1:N-1} \le v^*}\right) - \Pr(\max_j v_j > v^*) H\left(\delta_{v_{1:N-1} > v^*}\right)$$

$$< \Pr\left(\max_j v_j \le v^*\right) \Pr\left(v_i > v^*\right) \underbrace{\left[H\left(p\left(\cdot \mid v_i > v^*\right)\right) - \sum_{v > v^*} \frac{p(v)}{\Pr(v_i > v^*)} H(\delta_v)\right]}_{\text{gains from not learning } v_i \text{ fully when } v_i > v^* \text{ but } \max_j v_j \le v^*}$$

$$+ \Pr\left(\max_j v_j > v^*\right) \Pr\left(v_i \le v^*\right) \underbrace{\left[H\left(p\left(\cdot \mid v_i \le v^*\right)\right) - \sum_{v \le v^*} \frac{p(v)}{\Pr(v_i \le v^*)} H(\delta_v)\right]}_{\text{gains from not learning } v_i \text{ fully when } v_i \le v^* \text{ but } \max_j v_j > v^*}.$$

Intuitively, the LHS is the cost of $\Pi_{v^*}^{other}$ whereas the RHS is the reduction in information costs on self that buyer i achieves after acquiring signal $\Pi_{v^*}^{other}$ about others. The former is a fairly coarse partition—it only has two elements—and so does not grow as the prior gets uncertain. From the assumption of strong concavity, it just have to be above

$$\begin{aligned} &H(p_{1:N-1}) - \Pr(\max_j v_j \le v^*) H\left(\delta_{v_{1:N-1} \le v^*}\right) - \Pr(\max_j v_j > v^*) H\left(\delta_{v_{1:N-1} > v^*}\right) \\ &\quad \ge \frac{m}{2} \Pr(\max_j v_j \le v^*) \Pr(\max_j v_j > v^*) \left[\frac{\sum_{v \le v^*} (p_{1:N-1}(v))^2}{\Pr(\max_j v_j \le v^*)^2} + \frac{\sum_{v > v^*} (p_{1:N-1}(v))^2}{\Pr(\max_j v_j > v^*)^2}\right], \end{aligned}$$

but this lower bound goes to zero as $\sum_v (p(v))^2$ goes to zero.

On the contrary, the reduction in information costs on self grows as the prior becomes uncertain. To show this, we find a lower bound for the RHS of (1) using the fact that H is strongly concave. To apply the notion of strong concavity, we need to consider mixtures between two beliefs only. This can be done iteratively in the following way:

$$p\left(\cdot \mid v_i \le v^*\right) = \frac{p(v^1)}{\Pr(v_i \le v^*)} \delta_{v^1} + \left[1 - \frac{p(v^1)}{\Pr(v_i \le v^*)}\right] \delta_{v^1 < v_i \le v^*},$$

$$\text{where } \delta_{v^1 < v_i \le v^*} = \frac{p(v^2)}{\Pr(v^1 < v_i \le v^*)} \delta_{v^2} + \left[1 - \frac{p(v^2)}{\Pr(v^1 < v_i \le v^*)}\right] \delta_{v^2 < v_i \le v^*}$$

$$\text{and } \delta_{v^2 < v_i \le v^*} = \frac{p(v^3)}{\Pr(v^2 < v_i \le v^*)} \delta_{v^3} + \left[1 - \frac{p(v^3)}{\Pr(v^2 < v_i \le v^*)}\right] \delta_{v^3 < v_i \le v^*} \text{ , etc.}$$

Using the strong concavity of H iteratively, we then get

$$H\left(p\left(\cdot \mid v_j \le v^*\right)\right) - \sum_{v \le v^*} \frac{p(v)}{\Pr(v_i \le v^*)} H(\delta_v)$$

$$\geq \frac{m}{2} \frac{p(v^1)}{\Pr(v_i \leq v^*)} \left[1 - \frac{p(v^1)}{\Pr(v_i \leq v^*)}\right] ||\delta_{v^1} - \delta_{v^1 < v_i \leq v^*}||^2$$
$$+ \frac{m}{2} \left[1 - \frac{p(v^1)}{\Pr(v_i \leq v^*)}\right] \frac{p(v^2)}{\Pr(v^1 < v_i \leq v^*)} \left[1 - \frac{p(v^2)}{\Pr(v^1 < v_i \leq v^*)}\right] ||\delta_{v^2} - \delta_{v^2 < v_i \leq v^*}||^2$$
$$+ \dots$$

where

$$||\delta_{v^k} - \delta_{v^k < v_i \leq v^*}||^2 = 1 + \sum_{v^k < v_i \leq v^*} \left[\frac{p(v_i)}{\Pr(v^k < v_i \leq v^*)}\right]^2.$$

We hence get

$$H\left(p\left(\cdot \mid v_j \leq v^*\right)\right) - \sum_{v \leq v^*} \frac{p(v)}{\Pr(v_i \leq v^*)} H(\delta_v) \geq \frac{m}{2} \sum_{v \leq v^*} \frac{p(v)}{\Pr(v_i \leq v^*)} \frac{\Pr(v < v_i \leq v^*)}{\Pr(v \leq v_i \leq v^*)}$$
$$+ \frac{m}{2} \sum_{v \leq v^*} \frac{p(v)}{\Pr(v_i \leq v^*) \Pr(v < v_i \leq v^*) \Pr(v \leq v_i \leq v^*)} \sum_{v < v_i \leq v^*} [p(v_i)]^2.$$

Using the fact that $\frac{\Pr(v < v_i \leq v^*)}{\Pr(v \leq v_i \leq v^*)} = \frac{\Pr(v \leq v_i \leq v^*) - p(v)}{\Pr(v \leq v_i \leq v^*)}$, and collecting all $[p(v_i)]^2$ terms together, we get

(2) $$H\left(p\left(\cdot \mid v_j \leq v^*\right)\right) - \sum_{v \leq v^*} \frac{p(v)}{\Pr(v_i \leq v^*)} H(\delta_v) \geq \frac{m}{2} - \frac{m}{2} \frac{1}{\Pr(v_i \leq v^*)} \sum_{v \leq v^*} [p(v)]^2.$$

Similarly,

(2′) $$H\left(p\left(\cdot \mid v_j > v^*\right)\right) - \sum_{v > v^*} \frac{p(v)}{\Pr(v_i > v^*)} H(\delta_v) \geq \frac{m}{2} - \frac{m}{2} \frac{1}{\Pr(v_i > v^*)} \sum_{v > v^*} [p(v)]^2.$$

Note that these two lower bounds increase and tend to $m/2$ as the prior gets more uncertain.

We now find an upper bound for the cost of $\Pi_{v^*}^{other}$, that is for the LHS of (1). Note that

$$H(p_{1:N-1}) - \Pr(\max_j v_j \leq v^*) H\left(\delta_{v_{1:N-1} \leq v^*}\right) - \Pr(\max_j v_j > v^*) H\left(\delta_{v_{1:N-1} > v^*}\right)$$

$$= H(p_{1:N-1}) - \sum_v p_{1:N-1}(v) H\left(\delta_{v_{1:N-1}=v}\right)$$

$$- \Pr(\max_j v_j \leq v^*) \left[H\left(\delta_{v_{1:N-1} \leq v^*}\right) - \sum_{v \leq v^*} \frac{p_{1:N-1}(v)}{\Pr(\max_j v_j \leq v^*)} H(\delta_{v_{1:N-1}=v}) \right]$$

$$- \Pr(\max_j v_j > v^*) \left[H\left(\delta_{v_{1:N-1} > v^*}\right) - \sum_{v > v^*} \frac{p_{1:N-1}(v)}{\Pr(\max_j v_j > v^*)} H(\delta_{v_{1:N-1}=v}) \right].$$

Using the same steps as above, we can find bounds on the last two terms, such that

$$H(p_{1:N-1}) - \Pr(\max_j v_j \leq v^*) H\left(\delta_{v_{1:N-1} \leq v^*}\right) - \Pr(\max_j v_j > v^*) H\left(\delta_{v_{1:N-1} > v^*}\right)$$

$$\leq H(p_{1:N-1}) - \sum_v p_{1:N-1}(v) H(\delta_{v_{1:N-1}=v}) - \frac{m}{2}\left[1 - \sum_v [p_{1:N-1}(v)]^2\right].$$

Combining this bound on the LHS with the bounds (2) and (2′) on the RHS, condition (1) holds if

$$H(p_{1:N-1}) - \sum_v p_{1:N-1}(v)(\delta_{v_{1:N-1}=v}) - \frac{m}{2}\left[1 - \sum_v [p_{1:N-1}(v)]^2\right]$$

$$< \frac{m}{2}\left[\Pr\left(\max_j v_j \leq v^*\right) \Pr\left(v_i > v^*\right) + \Pr\left(\max_j v_j > v^*\right) \Pr\left(v_i \leq v^*\right)\right]$$

$$- \frac{m}{2}\left[\Pr\left(\max_j v_j > v^*\right) \sum_{v \leq v^*} (p(v))^2 + \Pr\left(\max_j v_j \leq v^*\right) \sum_{v > v^*} (p(v))^2\right].$$

Let $T \equiv \sum_v (p(v))^2$, which lies weakly below one. No value $v \in V$ can have prior probability greater than $\sqrt{T}$. Hence, for small T, it is possible to find v^* such that $\Pr(\max_j v_j \leq v^*)$ is close to 0.5. The RHS can then be arbitrarily close to $m/4$ for sufficiently small T. Hence, for T small enough, the RHS is strictly positive and increasing in m. Similarly, the LHS is positive but decreasing in m. For m high enough, the inequality must then hold strictly, establishing the claim. □

Lemma 2. *There exists $\overline{T}$ and $\overline{m}$ such that if $\sum_v [p(v)]^2 \leq \overline{T}$ and $m \geq \overline{m}$, then the following is true:*

$$c\left(\left\{\{v^1\}, \{v^2, \ldots, v^*\}, \{v : v > v^*\}\right\}, p_{1:N-1}\right) - c\left(\left\{\{v^1\}, \{v : v > v^1\}\right\}, p_{1:N-1}\right)$$

$$< \Pr\left(\max_j v_j > v^1\right) c\left(\{\{v\}_{v\in V}\}, p\right)$$

$$- \Pr\left(v^2 \leq \max_j v_j \leq v^*\right) c\left(\{\{v\}_{v\leq v^*}, \{v : v > v^*\}\}, p\right)$$

$$- \Pr\left(v^* < \max_j v_j\right) c\left(\{v : v \leq v^*\}, \{\{v\}_{v>v^*}\}, p\right),$$

for some $v^* > v^1$.

Lemma 2 says that the cost-efficient information structure has $\Pi^{other} \neq \{\{v^1\}, \{v : v > v^1\}\}$, such that buyers do not just learn whether their toughest opponent has the smallest possible valuation v^1 or not. This is important as such information structure would not lead to a revenue loss, as losing buyers would always learn their valuations fully.[33]

Proof. We rewrite the condition in terms of the measure of uncertainty H:

(3)

$$\Pr(\max_j v_j > v^1) H\left(\delta_{1:N-1>v^1}\right) - \Pr(v^1 < \max_j v_j \leq v^*) H\left(\delta_{v^1<1:N-1\leq v^*}\right)$$

$$- \Pr(\max_j v_j > v^*) H\left(\delta_{1:N-1>v^*}\right)$$

$$< \Pr\left(v^2 \leq \max_j v_j \leq v^*\right) \left[\Pr(v_i > v^*) H(\delta_{>v^*}) + \sum_{v\leq v^*} \Pr(v_i = v) H(\delta_v)\right]$$

$$+ \Pr\left(\max_j v_j > v^*\right) \left[\Pr(v_i \leq v^*) H(\delta_{\leq v^*}) + \sum_{v>v^*} \Pr(v_i = v) H(\delta_v)\right]$$

$$- \Pr\left(\max_j v_j > v^1\right) \sum_v \Pr(v_i = v) H(\delta_v).$$

Dividing everything by $\Pr(\max_j v_j > v^1)$, and using the fact that $H(\delta_{\leq v^*}) > \Pr(v_i = v^1) H(\delta_{v^1}) + \Pr(v^1 < v_i \leq v^*) H(\delta_{v^1<v_i\leq v^*})$, condition (3) holds if

$$H\left(\delta_{1:N-1>v^1}\right) - \Pr\left(\max_j v_j \in (v^1, v^*] \mid \max_j v_j > v^1\right) H\left(\delta_{v^1<1:N-1\leq v^*}\right)$$

[33] Indeed, after learning $\max_j v_j > v^1$, they would still choose the fully-revealing partition $\Pi^{self} = \{\{v\}_{v\in V}\}$.

$$
\begin{aligned}
&\qquad\qquad - \Pr\left(\max_j v_j > v^* \mid \max_j v_j > v^1\right) H\left(\delta_{1:N-1>v^*}\right) \\
&< \Pr\left(\max_j v_j \in (v^1, v^*] \;\middle|\; \max_j v_j > v^1\right)\left[\Pr(v_i > v^*)H(\delta_{>v^*}) - \sum_{v>v^*} \Pr(v_i = v)H(\delta_v)\right] \\
&\quad + \Pr\left(\max_j v_j > v^* \;\middle|\; \max_j v_j > v^1\right)\left[\Pr(v_i \leq v^*)H(\delta_{\leq v^*}) - \sum_{v \leq v^*} \Pr(v_i = v)H(\delta_v)\right]
\end{aligned}
$$

for some $v^* > v^1$. This is exactly the same condition as in the proof of Lemma 1, but for a redefined set of valuations $\widehat{V} = V \setminus v^1$. We then know from Lemma 1 that for $\sum_{v>v^1}\left(\frac{p(v)}{1-p(v^1)}\right)^2$ small enough and m large enough, the inequality holds for some $v^* \in \widehat{V}$, hence proving the claim. □

B.2 Proofs of Results of Section 3

B.2.1 Proofs of Proposition 1

Proof of Proposition 1. For Proposition 1, suppose that each buyer can only learn about himself. We look for a symmetric equilibrium in which, for λ small enough, buyers become fully informed. For this to be the case, buyers must choose the partition $\Pi_0 \equiv \{\{v_i\}_{v_i \in V}\}$ with a probability that tends to one as λ goes to zero. We construct a symmetric equilibrium σ that has such property.

Consider the following symmetric pure strategy profile:

- Each buyer chooses the finest partition $\Pi^{self} = \Pi_0$;
- Each buyer bids his valuation $\beta((\{v_i\}, u_i)) = v_i + u_i$.

It is a dominant strategy for a buyer who knows his value to bid his valuation for the good $\beta((\{v_i\}, u_i)) = v_i + u_i$. Hence we only need to check that buyers do not want to deviate to another information partition.

Let $K \equiv |V|$ and order the possible valuations for the good in increasing order: $V \equiv \{v^1, v^2, \ldots, v^K\}$ with $v^1 < v^2 < \cdots < v^K$. Any other information partition $\Pi^{self} \neq \Pi_0$ must bundle at least two possible valuations together. That is, there exist v^k and $v^{k'}$ that belong to the same element of the partition Π^{self}. Such bundling reduces a buyer's information costs by

$$\lambda\left(c\left(\Pi_0, p\right) - c\left(\Pi^{self}, p\right)\right) > 0.$$

We however show that such bundling must make buyer i strictly worse off in the auction, and so cannot be optimal for λ small enough.

First, we argue that Π^{self} cannot bundle two non-neighboring values v^k and $v^{k'}$ with $k' > k+1$. Recall that, by assumption, others follow their equilibrium strategy: they become fully informed of their values and bid truthfully. Hence, with probability $p_{1:N-1}(v^{k+1})$, the highest bid among i's competitors lies in $[v^{k+1} + \underline{u}, v^{k+1} + \overline{u}]$. At that bid, i wants to lose if his value is v^k and wants to win if his value is $v^{k'}$. Hence there is a strictly positive gain for him in distinguishing $v_i = v^k$ from $v_i = v^{k'}$, and for λ small enough, he must be doing so.

Second, we argue that Π^{self} cannot bundle two neighboring values v^k and v^{k+1} either. This is a bit more subtle, and would not be true absent the noise terms $(u_i)_i$ in buyers' valuations. Suppose a buyer bundles $\{v^k, v^{k+1}\}$. Upon learning that $v_i \in \{v^k, v^{k+1}\}$, it is (weakly) optimal for i to bid truthfully $\mathbb{E}[v_i|v_i \in \{v^k, v^{k+1}\}]+u_i$. Had buyer i chosen the fully revealing partition Π_0, he would have bid $v^k + u_i$ when $v_i = v^k$, and $v^{k+1} + u_i$ when $v_i = v^{k+1}$. If the highest bid among i's competitors lies above $v^{k+1} + u_i$ or below $v^k + u_i$, then this bundling does not change anything. Hence any difference in gross payoff between these two partitions must occur when i's toughest competitor, call him j^*, has a value $\max_j \nu_j \in (v^k + u_i, v^{k+1} + u_i)$. This requires either $v_{j^*} = v^k$ and $u_{j^*} > u_i$, or $v_{j^*} = v^{k+1}$ and $u_{j^*} < u_i$.

Focus on the first case, where $v_{j^*} = v^k$. Fix a realization of u_i, and consider what happens when $u_{j^*} > u_i$. There are two possible scenarios: either $u_{j^*} > u_i + \mathbb{E}[v_i|v_i \in \{v^k, v^{k+1}\}] - v^k$, in which case buyer i does not win when he bids $\mathbb{E}[v_i|v_i \in \{v^k, v^{k+1}\}] + u_i$. Learning to distinguish v^k from v^{k+1} allows the buyer to win the auction at the latter value, and hence induce a gain in gross payoff of

$$p(v^{k+1}) \left(v^{k+1} + u_i - v^k - u_{j^*}\right) > 0.$$

If $u_{j^*} \leq u_i + \mathbb{E}[v_i|v_i \in \{v^k, v^{k+1}\}] - v^k$, then i wins when he bids $\mathbb{E}[v_i|v_i \in \{v^k, v^{k+1}\}]$. Learning to distinguish v^k from v^{k+1} allows the buyer not to win the auction at the former value, and hence induces a gain in gross payoff of

$$p(v^k) \left(u_{j^*} - u_i\right) > 0.$$

Hence, as soon as $\Pr(u_{j^*} > u_i) > 0$, the expected gains from distinguishing these two values are strictly positive. This is the case whenever $\overline{u} > \underline{u}$, i.e. whenever the noise is not degenerate at zero.[34] For λ small enough, the cost of distinguishing these values must be strictly below the gains, and so the above strategy profile forms an equilibrium.

$\square$

B.2.2 Proofs of Proposition 2 and 3

Proof of Proposition 2. We first prove the existence of a symmetric equilibrium building on Corollary 5.3 in Reny (1999) (Step 1). We then argue that the proof extends for the existence of symmetric equilibria that are robust to trembles (Step 2).

Step 1. Without loss, we can restrict agents' possible bids to belong to $[0, \overline{\nu}]$ for some $\overline{\nu}$ that is greater than any possible realized valuation for the good. Agents' pure strategy sets are then compact Hausdorff spaces. Furthermore, their utility functions are bounded, measurable, and symmetric. Corollary 5.3 in Reny (1999) then states that there exists a symmetric mixed strategy equilibrium if (the mixed extension of) our game is better-reply secure along the diagonal. The crux of the proof is then to show that this condition of better-reply security holds in our setting.

Let $w_i(\sigma_i, \sigma_{-i})$ be buyer i's ex-ante expected utility given the strategy profile, that is

$$w_i(\sigma_i, \sigma_{-i}) \equiv \mathbb{E}_{\sigma_i,\sigma_{-i}} \left[U(\nu_i, \beta_i(\pi_i), \beta_{-i}(\pi_{-i})) - \lambda \left(c\left(\Pi_i^{other}, p_{1:N-1}\right) + c\left(\Pi_i^{self}\left(\pi_i^{other}\right), p\right)\right)\right].$$

Let $\sigma^* \in \Delta\Sigma$ denote a (potentially mixed) strategy. By symmetry, $w_i(\sigma^*, \ldots, \sigma^*)$ is independent of i. Let $w(\sigma^*) \equiv w_i(\sigma^*, \ldots, \sigma^*)$ be the diagonal payoff function, i.e., the payoff of an agent when all agents play the same strategy $\sigma^* \in \Delta\Sigma$. The game is diagonally better-reply secure if, whenever $(\sigma^*, w^*) \in \Delta\Sigma \times \mathbb{R}$ is in the closure of the graph of its diagonal payoff function and $(\sigma^*, \ldots, \sigma^*)$ is *not* an equilibrium, then some buyer i can secure a payoff strictly above w^* along the diagonal at $(\sigma^*, \ldots, \sigma^*)$.[35]

We show that our game satisfies this condition. Suppose that $(\sigma^*, \ldots, \sigma^*)$ is not an

[34] To see what goes wrong with the above argument when there is no noise—i.e. $\Pr(u_i = 0) = 1$—recall that buyers fully learn their values in the proposed equilibrium. Hence absent noise, other buyers' bids only take values in V. In particular, that means a buyer never faces a bid strictly in between v^k and v^{k+1}. There are thus no opportunity cost in bundling these values together, but only a strictly reduction in information costs. The equilibrium would then require buyers to mix between bundling neighboring values and not bundling them.

[35] Buyer i can secure a payoff strictly above w^* along the diagonal at $(\sigma^*, \ldots, \sigma^*)$ if there exists $\delta > 0$ and $\bar{\sigma}_i$, such that $w_i(\sigma', \ldots, \bar{\sigma}_i, \ldots, \sigma') \geq w^* + \delta$ for all σ' in some open neighborhood of σ^*.

equilibrium and let (σ^*, w^*) be an element of the closure of the graph of its diagonal payoff function. By definition, there exists a sequence of strategies $\sigma^{(n)}$ converging to σ^* such that $\lim w(\sigma^{(n)}) = w^*$. There are two cases: either w is continuous at σ^* or it is not. Discontinuities in the payoff can only arise from ties at the bidding stage. Hence the first case arises when, under σ^*, no tie occurs with strictly positive probability. In that case, $w(\sigma^*) = w^*$. Since σ^* is not an equilibrium, a buyer i has a deviation σ_i' that yields $w_i(\sigma_i', \sigma_{-i}^*) > w(\sigma^*) = w^*$. Furthermore, at this deviation $(\sigma_i', \sigma_{-i}^*)$, buyer i's payoff is continuous in others' strategy σ_{-i}. (If not, that means i ties with strictly positive probability under σ_i' and a marginally close strategy would make him strictly better off.) Hence i can secure a payoff within ε of $w_i(\sigma_i', \sigma_{-i}^*) > w^*$, as required by better-reply security.

Now consider the second case in which w is discontinuous at σ^*. This means relevant ties must occur with strictly positive probability at σ^*. Some of the buyers who tie with positive probability under σ^* must lose with positive probability at $\sigma^{(n)}$ for large enough n while they would be strictly better off winning. (Indeed, they cannot all be indifferent between winning and losing since a positive mass of them tie and they cannot all have the same expected value conditional on winning, since their u_is differ.) A buyer i can thus deviate to some σ_i' that breaks all ties in his favor, such that w_i is now continuous in σ_{-i} at $w_i(\sigma_i', \sigma_{-i}^{(n)}) > w_i(\sigma^{(n)})$ for large n. And so i can secure a payoff strictly above $\lim w_i(\sigma^{(n)}) = w^*$. Overall, the game is then better-reply secure and must admit a symmetric mixed-strategy equilibrium.

Step 2. We now argue that it must also admit a symmetric equilibrium that is robust to trembles. Consider a perturbed version of our game $G^{(k)}$ in which payoffs are perturbed with the $\varepsilon^{(k)}$, with $\lim \varepsilon^{(k)} = 0$.[36] For each k, the existence of a symmetric mixed-strategy equilibrium $(\sigma^{(k)})$ follows from the above argument.

We now need to show that the equilibrium of the perturbed game $\sigma^{(k)}$ converges to some feasible strategy $\sigma^* \in \Delta\Sigma$ as k grows large. This is true (at least for some subsequence) since the set of strategies $\Delta\Sigma$ to which $\sigma^{(k)}$ belongs is (sequentially) compact. So there exists a feasible strategy σ^* to which the equilibrium strategy $\sigma^{(k)}$ converges to as perturbations vanish.

[36] That is, agents have the same strategy space but the payoff associated with a strategy profile equals the expected payoffs if the agent's bid coincides with the one specified by his strategy with probability $1 - \varepsilon^{(k)}$ and is drawn from some exogenous continuous distribution F otherwise.

Finally, we have left to show that the strategy it converges to σ^* forms a symmetric equilibrium of the unperturbed game. By way of contradiction, suppose $(\sigma^*, \ldots, \sigma^*)$ is not an equilibrium. Then, by better-reply security, an agent i can secure a payoff strictly above $\lim w(\sigma^{(k)}) \equiv w^*$ using some strategy σ_i'. However σ_i' would then also be a profitable deviation against $\sigma_{-i} = \sigma^{(k)}$ for k large enough since payoffs are continuous in the perturbation $\varepsilon^{(k)}$, and so $\sigma^{(k)}$ cannot be an equilibrium of the perturbed game for k large enough. □

Proof of Proposition 3. By contradiction, suppose there exists a sequence of equilibria $\{\sigma_\lambda\}_\lambda$ such that $\lim_{\lambda \longrightarrow 0} \Pr\left(\Pi^{self} = \{\{v_i\}_{v_i \in V}\} \mid \sigma_\lambda\right) = 1$. That is, each buyer i might be mixing over partitions in equilibrium, but he must put a probability that tends to one as λ goes to zero on partition $\Pi^{self} = \Pi_0$.

We construct a profitable deviation. Consider an alternative strategy for buyer i, in which he first acquires information as to whether the maximum valuation among other bidders is above some threshold $v^{k^*} < v^K$, before learning about his own. That is, he chooses $\Pi^{other}_{k^*} = \{\{v^1, \ldots, v^{k^*}\}, \{v^{k^*+1}, \ldots, v^K\}\}$. Then, when he learns that his toughest competitor has a value above the threshold $\max_j v_j > v^{k^*}$, buyer i chooses to partition his set of valuations into $\Pi^{self}_{>k^*} \equiv \{\{v_i\}_{v_i > v^{k^*}}, \{v_i : v_i \leq v^{k^*}\}\}$. Intuitively, he does not learn to distinguish all the valuations below the threshold, as he most likely would not win at any of these. On the contrary, when he learns $\max_j v_j \leq v^{k^*}$ buyer i chooses the partition $\Pi^{self}_{\leq k^*} \equiv \{\{v_i\}_{v_i \leq v^{k^*}}, \{v_i : v_i > v^{k^*}\}\}$.

By Lemma 1, there exists k^* such that this alternative information strategy leads to strictly lower information cost than becoming fully informed about oneself:

$$c(\Pi_0, p) - \Pr(\max_j v_j > v^{k^*}) c(\Pi^{self}_{>k^*}, p) \\ - \Pr(\max_j v_j \leq v^{k^*}) c(\Pi^{self}_{\leq k^*}, p) - c(\Pi^{other}_{k^*}, \overline{p}_{N-1}) \equiv \Delta c > 0.$$

However, there is a potential opportunity cost of doing so if partitioning partially his set of valuations yields a lower gross payoff to i. (It has to yield a weakly lower payoff to i as information is valuable.) We now show that this opportunity cost is zero, and hence smaller than $\lambda \Delta c$.

Consider first what happens when i learns $\max_j v_j \leq v^{k^*}$. Since all agents converge to becoming fully informed about themselves by assumption, with a probability that tends to one they all bid at most $\max_j \nu_j \leq v^{k^*} + \overline{u} < v^{k^*+1} + \underline{u}$.

How does partition $\Pi^{self}_{\leq k^*}$ compare to Π_0? If buyer i has a value below the threshold v^{k^*}, then under both partitions he perfectly learns it and gets the same expected payoff. So these partitions can yield different payoff only when i has a value above the threshold. Under the former partition, i fails to distinguish his potential values, only learns $v_i \in \{v'_i \in V \mid v'_i > v^{k^*}\}$, and makes a bid in $[v^{k^*+1} + \underline{u}, v^K + \overline{u}]$. Under the latter partition, i learns his value v_i, and bids $v_i + u_i$. These partitions can then only yield different payoffs if one of i's competitors sometimes bids strictly above $v^{k^*+1} + \underline{u}$ despite his realized value being lower. Call that buyer j and that particular bid b^*. This can only be optimal for j if he failed to learn his value and chose a partition about himself that bundles his realized value $v_j \leq v^{k^*}$ with some other value(s) strictly above v^{k^*}. That is, buyer j bids b^* at an information set π_j such that $\min_{v_j \in \pi^{self}_j} v_j \leq v^{k^*}$ and $\max_{v_j \in \pi^{self}_j} v_j > v^{k^*}$.[37] Such bundling can be optimal for j only if he expects no non-vanishing cost from doing so, given others' equilibrium strategy. In particular, buyer j must believe that with a probability very close to one he will not face a bid in the interval of values he deems possible $[\min_{v_j \in \pi^{self}_j} v_j + \underline{u}, \max_{v_j \in \pi^{self}_j} v_j + \overline{u}]$. By assumption, buyers almost always learn and bid their valuations, and so buyer j must have learned that i's valuation is even higher than $\max_{v_j \in \pi^{self}_j} v_j$. If buyer i only learns that $v'_i > v^{k^*}$, he thus wants to make a bid that is weakly higher than any other bid he might face given equilibrium strategies, including b^*. He then gets the same gross payoff as if he had learned to distinguish these values, but at a lower information cost.

Overall, when i learns $\max_j v_j \leq v^{k^*}$, he knows for sure that if one of his competitors bids above $v^{k^*+1} + \underline{u}$, then he wants to match that bid. Doing so does not require learning the possible valuations he has that lie about v^{k^*}. Hence the two partitions about self, $\Pi^{self}_{\leq k^*}$ and Π_0, yield the exact same gross expected payoff.

Now consider what happens when i learns $\max_j v_j > v^{k^*}$. Buyer i then knows with certainty that $v_j > v^{k^*}$ for some j. Call him j^*. Given equilibrium strategies, that agent j^* converges to becoming fully informed about himself, so with a probability that tends to one makes a bid in $[v_{j^*} + \underline{u}, v_{j^*} + \overline{u}]$. How does partition $\Pi^{self}_{>k^*}$ compare to Π_0? Again, these partitions only differ when i has a value below the threshold. Under $\Pi^{self}_{>k^*}$, buyer i bundles all such values, and suppose that, in such a case, he bids sufficiently low so as to never win given others' equilibrium strategies. Under Π_0 he learns his value fully and bids $v_i + u_i$. Hence, for these two partitions to lead to different gross

[37]Note that our starting assumption simply requires buyers to become fully informed *with a probability close to one*. So this is possible, albeit with small probability.

expected payoffs, it has to be that with non-zero probability, the highest bid among i's competitors lies strictly below $v^{k^*} + \overline{u}$, and that by choosing such finer partition i sometimes wins.

Under what conditions can the highest-value buyer j^*, with value $\nu_{j^*} > v^{k^*} + \overline{u}$, bid strictly below $v^{k^*} + \overline{u}$ in equilibrium? For j^* to find such low bid optimal, it has to be that buyer j^* fails to learn his value: that is, he sometimes chooses a partition $\Pi_{j^*}^{self}$ that bundles $v_{j^*} > v^{k^*}$ with values below v^{k^*}. Denote that bundle by $\pi_{j^*}^{self}$. For such bundle to be optimal for j^*, it must be that j^* expects no non-vanishing loss from doing so. Since all buyers fully learn and bid their valuation with a probability close to one by assumption, this is only the case when j^* learned that $\max_{j \neq j^*} v_j < \min_{v'_{j^*} \in \pi_{j^*}^{self}} v'_{j^*}$. Hence j^*'s bid must lie above the valuations of all the other buyers, and none of them has any incentive to match his bid. In particular, buyer i never wants to win against j^*'s bid in such scenario, and choosing the finer partition cannot lead to strict gains. $\square$

B.2.3 Proof of Theorem 1

To prove Theorem 1, we find necessary conditions that must be satisfied by any information structure $\left(\Pi^{other}, \Pi^{self}\right) \in \mathcal{P} \times \mathcal{P}^{2^V}$ that has non-vanishing weight in some equilibrium σ_λ. Lemmas 3 to 5 show, in a succession of steps, that an equilibrium information structure $\left(\Pi^{other}, \Pi^{self}\right)$ must satisfy

$$(\star) \qquad \Pi^{self}(\pi^{other}) = \{\{v_i \mid v_i < \min_{v' \in \pi^{other}} v'\}, \{v_i\}_{v_i \in \pi^{other}}, \{v_i \mid v_i > \max_{v' \in \pi^{other}} v'\}\}$$

for all $\pi^{other} \in \Pi^{other}$. In words, after learning $\max_j v_j \in \pi^{other} \in \Pi^{other}$, the agent bundles all the values he can have that are below (resp. above) his toughest competitor's valuation for sure, as illustrated in Figure 2. Lemma 6 shows that an equilibrium information structure must also minimize total information costs. We know from Lemma 1 that this precludes $\Pi^{other} = \{V\}$, which completes the proof of Theorem 1.

Lemma 3. *There exists $\overline{\lambda} > 0$ such that, for all $\lambda \leq \overline{\lambda}$, there exists $\varepsilon(\lambda) > 0$ with $\lim_{\lambda \longrightarrow 0} \varepsilon(\lambda) = 0$ such that if an information structure has probability $\Pr\left(\Pi^{other}, \Pi^{self} \mid \sigma_\lambda\right) \geq \varepsilon(\lambda)$ in some equilibrium σ_λ, then it must have the following form: for all $\pi^{other} \in \Pi^{other}$,*

$$\{v_i\} \in \Pi^{self}\left(\pi^{other}\right) \quad \forall v_i \in \pi^{other}, v_i \neq \max_{v'_i \in \pi^{other}} v'_i.$$

In words, any information structure that has non-vanishing weight in equilibrium must satisfy the following condition: if agent i learns that his toughest competitor has value in some interval $\pi^{other} \equiv [v^{\underline{k}}, v^{\overline{k}}]$, then i at least learns to distinguish all the values he can have that lie in $[v^{\underline{k}}, v^{\overline{k}})$ as he knows competition will fall into that range.[38]

Proof of Lemma 3. Fix λ, and let $\left(\Pi^{other}, \Pi^{self}\right)$ be an information structure that is chosen with probability at least $\varepsilon > 0$ in some equilibrium. Take any $\pi^{other} \in \Pi^{other}$, and let $v^{\underline{k}} \equiv \min_{v'_i \in \pi^{other}} v'_i$ and $v^{\overline{k}} \equiv \max_{v'_i \in \pi^{other}} v'_i$. That is, upon learning $\max_j v_j \in \pi^{other}$, agent i knows that his toughest competitor has $v_j \in [v^{\underline{k}}, v^{\overline{k}}]$.

We prove that, after learning $\max_j v_j \in \pi^{other}$, the equilibrium partition that an agent chooses about himself $\Pi^{self}\left(\pi^{other}\right)$ cannot bundle a value $v^* \in \pi^{other}$ with some other value weakly below $v^{\overline{k}}$.[39] By contradiction, suppose that this is not true: after learning that $\max_j v_j \in \pi^{other}$, an agent chooses a partition that bundles some $v^* \in \pi^{other}$ with another that is weakly below $v^{\overline{k}}$. That bundle can be composed of only these two values, or can have other values in it too. Let $\underline{v}$ (resp, $\overline{v}$) be the lowest (resp, highest) element in that bundle, and denote that bundle by $\pi^{self}_{\underline{v},\overline{v}}$. Note that, without loss, $v^* > \underline{v}$. There are indeed two cases: either the bundle is only composed of values in π^{other}, in which case v^* can be any value in the bundle but the smallest one, or it isn't, in which case $\underline{v} < \min_{v'_i \in \pi^{other}} v'_i \leq v^*$.

Let $j^* \in \arg\max_j v_j$ be (any one of) agent i's toughest competitor(s).[40] At information set $\left(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i\right)$ agent i knows that his toughest competitor has $\max_j v_j \in \pi^{other}$ and that his own value v_i might also belong to that set, since $\pi^{other} \cap \pi^{self}_{\underline{v},\overline{v}} \neq \emptyset$. By the symmetry of the equilibrium, he hence knows that, with strictly positive probability, j^*'s information set is also $\pi_{j^*} = \left(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_{j^*}\right)$ for some u_{j^*}. With non-trivial probability, this is the highest bid that agent i faces. Indeed, it is for instance the case when all other agents choose the same information structure as i (with happens with probability ε^{N-1}) and have all value $v_j = v^*$. For his equilibrium bid to be optimal, agent i must then be indifferent between losing the auction and winning *at his equilibrium bid*

[38]Showing that buyer i also wants to distinguish valuation $v_i = v^{\overline{k}}$ is a bit more subtle, and we handle that case in Lemma 4.

[39]This is enough to prove the result. The other alternative is that v^* is bundled with values weakly above $v^{\overline{k}}$. If $v^* = v^{\overline{k}}$, this is irrelevant as Lemma 3 makes no claim about $v^{\overline{k}}$. The only relevant case is then $v^* < v^{\overline{k}}$. But if v^* is bundled with values above $v^{\overline{k}}$ then it must be bundled with $v^{\overline{k}}$ as well. However that implies $v^{\overline{k}}$ is bundled with a value v^* weakly below itself, and this is precisely what we are proving cannot happen.

[40]Note that j^* might not be the highest-value bidder if there exists another agent j with $v_j = v_{j^*}$ and $u_j > u_{j^*}$, but this is irrelevant for the following argument.

(i.e., winning at a tie, such that he pays his equilibrium bid):

$$\begin{aligned}\beta\left(\pi^{other},\pi^{self}_{\underline{v},\overline{v}},u_i\right) &= \mathbb{E}\left[v_i \;\middle|\; \pi_i=\left(\pi^{other},\pi^{self}_{\underline{v},\overline{v}},u_i\right),\max_j \beta(\pi_j)=\beta(\pi_i)\right]+u_i \\ &= \sum_{v_i=\underline{v}}^{\overline{v}} v_i \Pr\left[v_i \;\middle|\; \pi_i=\left(\pi^{other},\pi^{self}_{\underline{v},\overline{v}},u_i\right),\max_j \beta(\pi_j)=\beta(\pi_i)\right]+u_i.\end{aligned}$$

If not, then agent i would have an incentive to marginally increase or decrease his bid, depending on whether or not he wants to win the auction at that price. Note that $\beta(\pi^{other},\pi^{self}_{\underline{v},\overline{v}},u_i)\in[\underline{v}+u_i,\overline{v}+u_i]$ since $\underline{v}$ (resp. $\overline{v}$) is the lowest (resp. highest) value of v_i that agent i deems possible at information set $\left(\pi^{other},\pi^{self}_{\underline{v},\overline{v}},u_i\right)$.

We now show that agent i has a strict, non-vanishing incentive to learn to distinguish values $v_i=\underline{v}$ and $v_i=\overline{v}$, and will hence do so for small enough information costs. We know that, with a strictly positive probability that is increasing in ε and independent of λ, the highest bid made by i's competitors when i's information set is $\pi_i=(\pi^{other},\pi^{self}_{\underline{v},\overline{v}},u_i)$ lies in $[\underline{v}+u_j,\overline{v}+u_j]$ for some u_j. There are then strictly positive gains from unbundling values $v_i=\overline{v}$ and $v_i=\underline{v}$ since i wants to win against such bid when $v_i=\overline{v}$ and $u_i>u_{j^*}$ but wants to lose when $v_i=\underline{v}$ and $u_i<u_{j^*}$. (This is the same argument as in the proof of Proposition 1.)

There is a strictly positive cost $\lambda\Delta c$ associated with unbundling these values, as it requires choosing a finer partition and splitting $\pi^{self}_{\underline{v},\overline{v}}$ into at least two elements. However, for λ small enough, there exists $\varepsilon(\lambda)$ such that, if $\varepsilon\geq\varepsilon(\lambda)$ then the value of distinguishing these values more than compensates the cost. Hence there exists $\varepsilon(\lambda)$ such that, if an information structure has probability $\Pr\left(\Pi^{other},\Pi^{self}\mid\sigma_\lambda\right)\geq\varepsilon(\lambda)$, then it cannot make such a bundle. Furthermore, since the cost of unbundling these values goes to zero when λ goes to zero, then $\varepsilon(\lambda)$ must go to zero as well. □

Lemma 4. *There exists $\overline{\lambda}>0$ such that, for all $\lambda\leq\overline{\lambda}$, there exists $\varepsilon(\lambda)>0$ with $\lim_{\lambda\longrightarrow 0}\varepsilon(\lambda)=0$ such that if an information structure has probability $\Pr\left(\Pi^{other},\Pi^{self}\mid\sigma_\lambda\right)\geq\varepsilon(\lambda)$ in some equilibrium σ_λ, then it must have the following form:*

$$\left\{\max_{v_i'\in\pi^{other}} v_i'\right\}\in\Pi^{self}\left(\pi^{other}\right)\quad \textit{for all } \pi^{other}\in\Pi^{other}.$$

Proof of Lemma 4. Suppose not: there exists $\varepsilon>0$ and a sequence of equilibria $\{\sigma_\lambda\}_\lambda$

such that $\Pr\left(\Pi^{other}, \Pi^{self} \mid \sigma_\lambda\right) \geq \varepsilon$ for all λ and $\left\{\max_{v_i' \in \pi^{other}} v_i'\right\} \notin \Pi^{self}\left(\pi^{other}\right)$ for some $\pi^{other} \in \Pi^{other}$. Let $\overline{v} = \max_{v_i' \in \pi^{other}} v_i'$ and $\underline{v} = \min_{v_i' \in \pi^{other}} v_i'$. In words, after learning that $\max_j v_j \in [\underline{v}, \overline{v}]$, agent i does not learn to distinguish $v_i = \overline{v}$.

We know from Lemma 4 that $\Pi^{self}\left(\pi^{other}\right)$ cannot bundle $v_i = \overline{v}$ with lower values since an agent must learn to distinguish all values $v_i \in \pi^{other} \setminus \overline{v}$. Hence $\overline{v}$ must be bundled with even greater values, and denote this bundle by $\pi^{self}_{\geq \overline{v}}$. Let $\hat{v}$ denote the maximum valuation in $\pi^{self}_{\geq \overline{v}}$, which by definition satisfies $\hat{v} > \overline{v}$.

Step 1. We first show that, at information set $\pi_i = (\pi^{other}, \pi^{self}_{\geq \overline{v}}, u_i)$, an agent must bid arbitrarily close to $\overline{v} + u_i$ for sufficiently small λ. That is, for any $\eta > 0$ there exists $\bar{\lambda}$ such that, for all $\lambda \leq \bar{\lambda}$, $|\beta(\pi^{other}, \pi^{self}_{\geq \overline{v}}, u_i) - \overline{v} - u_i| \leq \eta$. Note that, at this information set, agent i knows that $\max_j v_j \in [\underline{v}, \overline{v}]$ and $v_i \in [\overline{v}, \hat{v}]$. Since the equilibrium is symmetric, he also knows that, with non-vanishing probability, his toughest opponent has the same information set and makes the same equilibrium bid as i. Indeed, this happens whenever $v_i = \overline{v}$, $\max_j v_j = \overline{v}$, and i's toughest opponent(s) chooses the same information structure as i. Note first that an equilibrium bid at $\pi_i = (\pi^{other}, \pi^{self}_{\geq \overline{v}}, u_i)$ cannot be bounded below the lowest possible valuation at that information set $\overline{v} + u_i$. Indeed, since an agent sometimes ties at that bid, he could slightly increase his bid and make a strict gain. More importantly, $\beta(\pi^{other}, \pi^{self}_{\geq \overline{v}}, u_i)$ cannot be bounded above $\overline{v} + u_i$ either. If it were the case, then the agent would have an incentive to learn whether $v_i = \overline{v}$ or $v_i > \overline{v}$, since he would not want to win against $\beta(\pi^{other}, \pi^{self}_{\geq \overline{v}}, u_i)$ in the former case. For sufficiently small information cost λ, he would do so.

Step 2. We then show that such bid at information set $\pi_i = (\pi^{other}, \pi^{self}_{\geq \overline{v}}, u_i)$ cannot be part of an equilibrium. If it were, then an agent i would make such a bid with probability at least ε in all states of the world consistent with information set π_i. In particular, he would make such a bid in states where his toughest opponent has value $\max_j v_j = \overline{v}$ and he has value $v_i = \hat{v}$. But then his toughest opponent would have a strictly positive, non-vanishing incentive to learn that his valuation is $v_j = \overline{v}$, as he then wants to outbid i if $u_j > u_i$ and to lose if $u_j < u_i$. He would do so in equilibrium, and agent i would have a strict incentive to deviate and learn that his value is $v_i = \hat{v}$. $\square$

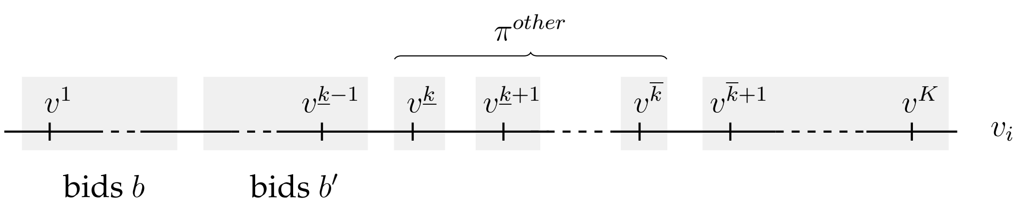
Lemma 5. *There exists $\overline{\lambda} > 0$ such that, for all $\lambda \leq \overline{\lambda}$, there exists $\varepsilon(\lambda) > 0$ with $\lim_{\lambda \longrightarrow 0} \varepsilon(\lambda) = 0$ such that if an information structure has probability $\Pr\left(\Pi^{other}, \Pi^{self} \mid \sigma_\lambda\right) \geq \varepsilon(\lambda)$ in some* *equilibrium σ_λ, then it must have the following form: for all $\pi^{other} \in \Pi^{other}$,*

$$\left\{v_i \mid v_i < \min_{v_i' \in \pi^{other}} v_i'\right\} \in \Pi^{self}\left(\pi^{other}\right) \quad \textit{and} \quad \left\{v_i \mid v_i > \max_{v_i' \in \pi^{other}} v_i'\right\} \in \Pi^{self}\left(\pi^{other}\right).$$

Proof of Lemma 5. Fix λ and let $\left(\Pi^{other}, \Pi^{self}\right)$ be an information partition that is chosen with probability at least ε in equilibrium σ_λ. Take any $\pi^{other} \in \Pi^{other}$, and let $v^{\underline{k}} \equiv \min_{v_i' \in \pi^{other}} v_i'$ and $v^{\overline{k}} \equiv \max_{v_i' \in \pi^{other}} v_i'$. That is, upon learning $\max_j v_j \in \pi^{other}$, buyer i knows that his toughest competitor has $v_j \in [v^{\underline{k}}, v^{\overline{k}}]$.

Step 1. We prove that if i learns $\max_j v_j \in \pi^{other}$, then i chooses a signal about himself that bundles all the values he might have that lie for sure below his toughest competitor's valuation: $\{v_i \mid v_i < v^{\underline{k}}\} \in \Pi^{self}(\pi^{other})$. Denote by $j^* \in \arg\max_j v_j$ (one of) i's toughest competitors.

By contradiction, suppose that the claim is not true, and that in equilibrium buyer i partitions all these values into at least two elements:



Since choosing such finer partition is costly, it must lead him to make better decision-making: so he must make (at least two) different bids b and b' depending on what he learned, and these two bids lead to different outcomes. Hence it must be that, with some strictly positive probability, the highest bid faced by buyer i lies in between these two bids $\Pr(b \leq \max_j \beta(\pi_j) \leq b' \mid \max_j v_j \in \pi^{other}, v_i < v^{\underline{k}}) > 0$, as otherwise these two bids would be completely equivalent.

This implies that buyer j^*, whom we know has a value $v_{j^*} \in \pi^{other}$, must sometimes make a bid below $b' \leq v^{\underline{k}-1} + u_i < v^{\underline{k}} + \underline{u}$.[41] For j^* to make such a low bid in equilibrium, he must fail to learn that he has a high value and bundle his high value with lower ones. Let π_{j^*} be the information set at which j^* acts in such a way, with $\beta(\pi_{j^*}) \leq b'$.

[41]The first inequality comes from the fact that in a tremble-robust equilibrium a buyer can never make a bid that lies *outside* the set of valuations he deems feasible. When i makes bid b', he knows that his value is at most $\nu_i \leq v^{\underline{k}-1} + u_i$. The second inequality holds by construction since the noise terms are smaller than the size of the grid defined by V.

We furthermore know that j^* sometimes has this information set when he is a highest-valuation buyer and his value $v_{j^*} \in \pi^{other}$. Hence $\max_{v \in \pi^{self}_{j^*}} \geq v^{\underline{k}}$.

Failing to learn his valuation is however costly for j^* as it leads him to sometimes lose the auction against i's bid, despite his value being higher than i's winning bid. Note that j^* cannot lose against i's bid with non-vanishing probability, as otherwise it would be profitable for him to learn his valuation and bid it for λ small enough. For such bundle to be optimal for j^*, he must then expect to face a bid weakly above $v_{j^*} + \overline{u}$ with a probability that tends to one as λ goes to zero. If not, then learning to distinguish his high value from lower ones would lead to strictly positive, non-vanishing gains, and for λ small enough he would do so. Hence there must be another buyer, call him k, who bids weakly above $v_{j^*} + \overline{u}$ with non-vanishing probability in those states. Since j^* is a highest-valuation buyer, it must be that $\nu_k \leq \nu_{j^*} < v_{j^*} + \overline{u}$ almost surely, and hence that buyer k bids strictly above his valuation. By the symmetry of the equilibrium, buyer k must however tie at his equilibrium bid with non-vanishing probability.[42] He then incurs a strict, non-vanishing loss when that occurs, and so such a high bid by k cannot be sustained in equilibrium.

Step 2. Finally, we prove that if i learns $\max_j v_j \in \pi^{other}$, then i chooses a signal about himself that bundles all the values he might have that he knows for sure lie above his toughest competitor's valuation: $\left\{v_i \mid v_i > v^{\overline{k}}\right\} \in \Pi^{self}\left(\pi^{other}\right)$.

By contradiction, suppose this is not the case. Following a similar logic as for Step 1, agent i can only find it worthwhile to learn to distinguish some of the values $v_i > v^{\overline{k}}$ if he sometimes faces a bid in that interval, and sometimes loses at that bid. That means that with positive probability, one of i's competitors, all of whom have a value at most $v^{\overline{k}} + \overline{u}$, makes a bid strictly above this and wins with non-vanishing probability at that bid. However, if they win they must be paying a price weakly higher than i's bid, and i's bid must lie above $v^{\overline{k}} + \overline{u}$. Hence making such a bid leads that agent to have a strictly negative payoff, while he could ensure himself zero by making a lower bid. This cannot be optimal, and hence cannot be part of an equilibrium. $\square$

Lemma 6. *There exists $\overline{\lambda} > 0$ such that, for all $\lambda \leq \overline{\lambda}$, there exists $\varepsilon(\lambda) > 0$ with $\lim_{\lambda \longrightarrow 0} \varepsilon(\lambda) = 0$ such that if an information structure has probability $\Pr\left(\Pi^{other}, \Pi^{self} \mid \sigma_\lambda\right) \geq \varepsilon(\lambda)$ in some*

[42] If buyer k makes such a bid with non-vanishing probability then others must do as well.

equilibrium σ_λ, then it must be cost-minimizing, that is:

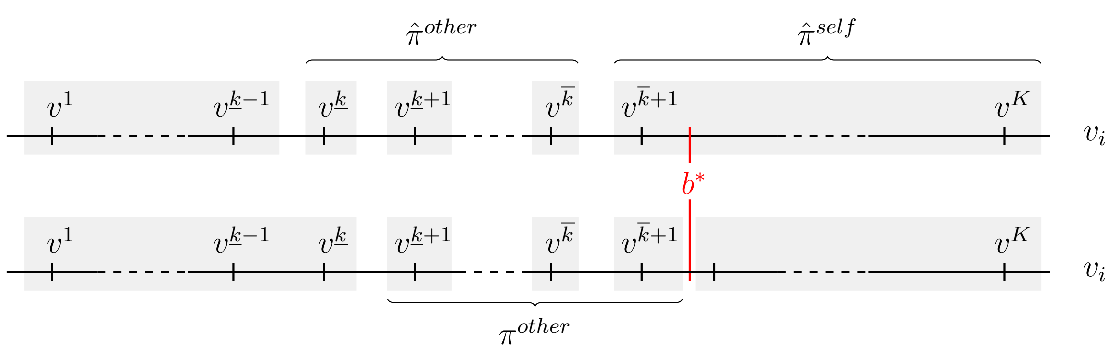
$$\left(\Pi^{other}, \Pi^{self}\right) \in \arg \min_{\widehat{\Pi}^{other}, \widehat{\Pi}^{self}} \quad c\left(\widehat{\Pi}^{other}, p_{1:N-1}\right) + \mathbb{E}_{\hat{\pi}^{other}}\left[c\left(\widehat{\Pi}^{self}(\hat{\pi}^{other}), p\right)\right]$$

$$s.t. \quad \left(\widehat{\Pi}^{other}, \widehat{\Pi}^{self}\right) \textit{ satisfies } (\star).$$

Proof of Lemma 6. Suppose not, such that an equilibrium puts non-vanishing probability $\varepsilon > 0$ on some information structure $\left(\Pi^{other}, \Pi^{self}\right)$ that is not cost-minimizing. That is, there exists another information structure $\left(\widehat{\Pi}^{other}, \widehat{\Pi}^{self}\right)$ satisfying ($\star$) that is strictly cheaper than $\left(\Pi^{other}, \Pi^{self}\right)$. Let $\lambda \Delta c$ be the difference in information costs between these two information structures.

Take the point of view of some agent i. For $\left(\widehat{\Pi}^{other}, \widehat{\Pi}^{self}\right)$ not to be a profitable deviation from $\left(\Pi^{other}, \Pi^{self}\right)$, it has to be that, under the former, agent i sometimes gets a strictly lower gross payoff at the auction stage. Since $\left(\widehat{\Pi}^{other}, \widehat{\Pi}^{self}\right)$ satisfies ($\star$), this can only happen when i fails to learn his valuation fully under $\left(\widehat{\Pi}^{other}, \widehat{\Pi}^{self}\right)$. That is, it can only happen when either $\hat{\pi}_i^{self} = \{v_i \mid v_i > \max_{v'_j \in \hat{\pi}_i^{other}} v'_j\}$ or $\hat{\pi}_i^{self} = \{v_i \mid v_i < \min_{v'_j \in \hat{\pi}_i^{other}} v'_j\}$.

Consider the first case, which we illustrate in the figure below. For agent i not to get his full-information optimal payoff at $(\hat{\pi}_i^{other}, \hat{\pi}_i^{self} = \{v_i \mid v_i > \max_{v'_j \in \hat{\pi}_i^{other}} v'_j\})$, it has to be that the highest bid i faces at this information set sometimes falls strictly within $[\min_{v_i \in \hat{\pi}_i^{self}} v_i + \underline{u}, \max_{v_i \in \hat{\pi}_i^{self}} v_i + \overline{u}]$. Call b^* such a bid and let j^* be an opponent



that submits it. For agent i to be strictly better off under $\left(\Pi^{other}, \Pi^{self}\right)$ because of it, it has to be that, under $\left(\Pi^{other}, \Pi^{self}\right)$, he better discriminates whether he should win against b^*. That is, for some values of $v_i \in \hat{\pi}_i^{self}$ with $v_i < b^*$, i learns his value and bids below b^*. (In the figure, this corresponds to i learning to distinguish $v_i = v^{\bar{k}+1}$

from $v_i > v^{\bar{k}+1}$ under $\left(\Pi^{other}, \Pi^{self}\right)$, as illustrated in the bottom partition.) However, that means agent j^* is making a strict loss winning against that bid, since we know $v_{j^*} \leq \max_{v_j \in \hat{\pi}_i^{other}} v_j < v_i$. Since information structure $\left(\Pi^{other}, \Pi^{self}\right)$ has non-vanishing probability, agent j^* is making a non-vanishing loss, and for small enough λ, he must find it profitable to learn enough so as to avoid this costly mistake.

The reasoning for the second case is very similar. □

Wrapping up, if an information structure has non-vanishing weight in some equilibrium, then it must satisfy ($\star$) and be cost-minimizing. The information structure under which agents acquire no information about others and fully learn their own valuation does satisfy ($\star$). However, it is not cost-minimizing (Lemma 1). Hence if an information structure has non-vanishing weight, it must involve acquire some information about $\max_j v_j$.

B.3 Proof of Results of Section 4

B.3.1 Proof of Theorem 2

We know from Theorem 1 that, for λ small enough, the only information structures that have non-trivial probability must satisfy:

$$\Pi^{self}(\pi^{other}) = \left\{ \left\{ v_i \mid v_i < \min_{v \in \pi^{other}} v \right\}, \left\{ v_i \right\}_{v_i \in \pi^{other}}, \left\{ v_i \mid v_i > \max_{v \in \pi^{other}} v \right\} \right\} \; \forall \pi^{other} \in \Pi^{other},$$

and must minimize total information costs. A direct implication is that buyers do acquire some information about the competition $\Pi^{other} \neq \{V\}$ (Lemma 1), and as a result sometimes fail to learn their valuations precisely. The proof of Theorem 2 leverages this to show that, in any equilibrium, expected revenue remains bounded away from the expected second-highest valuation even as the cost parameter λ goes to zero. We first show that, when losing buyers fail to learn their valuations, they bid their expected valuations given their information sets (Step 1). We then show that such behavior reduces the second-highest bid in expectation (Step 2).

Step 1. We show that, when a buyer fails to learn his valuations fully but only learns that he does not have the highest one—i.e., when $\pi_i^{self} = \{v_i \mid v_i < \min_{v'_i \in \pi_i^{other}} v'_i\}$,—then he bids his expected valuations given his information set. Denote that bundle by

$\pi_<^{self}$. We already know that, in any tremble-robust equilibrium, a buyer cannot make a bid that lies outside the interval of values he deems possible:

$$\beta(\pi^{other}, \pi_<^{self}, u_i) \in \left[\min_{v_i' \in \pi_<^{self}} v_i' + u_i, \max_{v_i' \in \pi_<^{self}} v_i' + u_i\right].$$

Furthermore, at that information set, a buyer must lose the auction with probability one in equilibrium. Suppose not, and let $j^* \in \arg\max_j v_j$ be (one of) i's toughest competitors. We know buyer i's bid must lie below j^*'s value since $\beta(\pi^{other}, \pi_<^{self}, u_i) < \min_{v_i' \in \pi^{other}} v_i' + \underline{u} \leq \nu_{j^*}$. We also know that buyer i chooses this information structure with non-vanishing probability. For i to win the auction at that information set, buyer j^* must fail to learn his valuation and bundle his high value in π^{other} with some lower ones. However such information structure cannot be optimal for j^* when λ is small enough as it leads him to lose against i's bid with non-vanishing probability. Hence it must be that at information set $\pi_i = (\pi^{other}, \pi_<^{self}, u_i)$, i always loses the auction in equilibrium.

Buyer i's equilibrium bid is then disciplined by the trembling-hand-like refinement that we impose. In particular, buyer i's bid must be optimal given that each of his competitors trembles with vanishing probability. Hence i can only win the auction when j^* trembles. In that scenario, none of i's competitors' bids reveal any information about i's value: all buyers $j \neq j^*$ learned about their toughest competitors, which is j^*, and j^* is trembling so his bid is drawn at random. Given that bidding truthfully is a weakly dominant strategy in a SPA, and that buyer i faces a distribution of bids that has full support given j^*'s tremble, he has a strict incentive to bid his expected valuation for the good given his information set:

$$\beta\left(\pi^{other}, \pi_<^{self}, u_i\right) = \mathbb{E}\left[v_i \,\middle|\, v_i < \min_{v_i' \in \pi^{other}} v_i'\right] + u_i.$$

Step 2. We show that there exist $L > 0$ and $\overline{\lambda} > 0$ such that, for all $\lambda \leq \overline{\lambda}$, the expected second-highest bid is lower than $\mathbb{E}[\nu_{(2)}] - L$ in any equilibrium.[43] Take any equilibrium, and denote by $q(\lambda)$ the probability that a buyer chooses an information structure satisfying ($\star$). We know from Theorem 1 that $\lim_{\lambda \longrightarrow 0} q(\lambda) = 1$. We focus on the case where all buyers choose such an information structure—the other case has vanishing

[43] $\nu_{(2)}$ denotes the second highest value for every realization of $(\nu_i)_i$.

probability, and induces a revenue that is bounded above by the highest possible valuation. We first show that, *given any realized second-highest bid* $b_{(2)}$, revenue (i.e., the realized second-highest bid) must lie weakly below $\mathbb{E}[\nu_{(2)} \mid b_{(2)}]$. We then show that, with strictly positive probability, it is bounded strictly below $\mathbb{E}[\nu_{(2)} \mid b_{(2)}] - L$ for some $L > 0$.

Note that the highest-valuation buyer wins with probability one if all buyers choose an information structure satisfying ($\star$). Let i_1 be the highest-valuation buyer and $\nu_{(1)}$ his realized valuation. (Ties have probability zero as the distribution of u_i is continuous.) There are two cases: either the price is set by a buyer who learned his valuation fully, or it isn't. The first case is direct: since the second-highest bidder (call him j^*) learned his value, he must have bid truthfully, and revenue then equals $b_{(2)} = \nu_{j^*} \leq \nu_{(2)}$. In the second case, the second-highest bidder failed to learn his value, which means that $\pi_{j^*}^{self} = \{v_j \mid v_j \leq \overline{v}\}$ for some $\overline{v} < v_{i_1}$. If j^* wins at such bid, then all other bidders $j \neq j^*, i_1$ must also have had values $v_j \leq \overline{v}$.[44] We know from Step 1 that j^* must have bid $\beta(\pi_{j^*}) = \mathbb{E}[\nu_{j^*} \mid v_{j^*} \leq \overline{v}]$, which always lies weakly below $\mathbb{E}[\nu_{(2)} \mid b_{(2)} = \beta(\pi_{j^*})] = \mathbb{E}[\nu_{(2)} \mid v_{i_1} > \overline{v}, v_j \leq \overline{v}\ \forall j \neq i_1]$.[45] Thus, when all agents choose information structures satisfying ($\star$), $b_{(2)} \leq \mathbb{E}[\nu_{(2)} \mid b_{(2)}]$, and expected revenue is weakly below the expected second-highest valuation.

We now prove that, with strictly positive non-vanishing probability, the second-highest bid $b_{(2)}$ is bounded strictly below the expected second-highest valuation given $b_{(2)}$. In particular, we show that this is the case when all buyers choose the same information structure and the gap between the highest and second-highest valuations is large enough. Let $V = \{v^1, v^2, \ldots, v^K\}$ with $v^{k+1} > v^k$. Take any information structure $(\Pi^{other}, \Pi^{self})$ that has non-vanishing weight in equilibrium. Let $\underline{v} = \min\{v \mid \exists \pi^{other} \in \Pi^{other} \text{ s.t. } v \in \pi^{other}, v^K \in \pi^{other}\}$ denote the smallest valuation that Π^{other} bundles with v^K. We know from Theorem 1 and Lemma 1 that $\Pi^{other} \neq \{V\}$. Similarly, we know from 2 that $\Pi^{other} \neq \{\{v^1\}, \{v^2, \ldots, v^K\}\}$. Hence $\underline{v} > v^2$. Consider what happens when the highest-valuation bidder has value $v_{i_1} \geq \underline{v}$ while all others $j \neq i_1$ have value $v_j < \underline{v}$. When all buyers choose information structure $(\Pi^{other}, \Pi^{self})$, all buyers $j \neq i_1$ must learn $\max_{j \neq i} v_j \in [\underline{v}, v^K]$ since $v_{i_1} \geq \underline{v}$. Furthermore, all buyers but i_1 must fail to learn their valuations precisely: $\pi_j^{self} = \{v_j \mid v_j < \underline{v}\}$ with $|\pi_j^{self}| \geq 2$. The second-highest bid

[44]If not, they would have either learned their values and outbid j^*, or just learned that $v_j \leq \overline{v}'$ for some $\overline{v}' > \overline{v}$, which would also have led them to outbid j^*.

[45]Indeed, the latter is the expectation of the highest value among the $N - 1 \geq 2$ losing bidders, which is weakly above the expected value of one specific losing bidder j^*.

then equals $\mathbb{E}[v_j \mid v_j < \underline{v}] + \max_{j \neq i_1} u_j$.

Overall, we get

$$\mathbb{E}\left[\text{ equilibrium revenue } |\sigma_\lambda\right] - \mathbb{E}[\nu_{(2)}] \leq \left(1 - [q(\lambda)]^N\right)\left(v^K - \mathbb{E}[\nu_{(2)}]\right)$$
$$+ [q(\lambda)]^N \Pr\left(\Pi^{other}, \Pi^{self} \mid \sigma_\lambda\right)^N \Pr\left(v_{i_1} \geq \underline{v}, v_j < \underline{v}\ \forall j \neq i_1\right) \times$$
$$\left(\mathbb{E}[v_i \mid v_i < \underline{v}] + \max_{j \neq i_1} u_j - \mathbb{E}[v_{(2)} \mid v_{(2)} < \underline{v}, v_{(1)} \geq \underline{v}] - u_{i_2}\right).$$

We give an upper bound for the second term. Note that $\mathbb{E}[v_i \mid v_i < \underline{v}] - \mathbb{E}[v_{(2)} \mid v_{(2)} < \underline{v}, v_{(1)} \geq \underline{v}]$ is strictly negative. Indeed, it compares the expected value of a buyer conditional on it being lower than some bound $\mathbb{E}[v_i \mid v_i < \underline{v}]$ to the expected *second-highest* value conditional on it being lower than that same bound and the highest-value being higher than this bound $\mathbb{E}[v_{(2)} \mid v_{(2)} < \underline{v}, v_{(1)} \geq \underline{v}]$. Hence the latter is just the expected value of the best of these $N-1$ draws, simply truncating the distribution at the bound as we know that all these $N-1$ draws lie below it. Since there are $N \geq 3$ buyers, the latter is strictly positive whenever there is some variance in the distribution of $v_i < \underline{v}$. This is the case as $|\{v_i \mid v_i < \underline{v}\}| \geq 2$. Hence

$$\mathbb{E}[v_{(2)} \mid v_{(2)} < \underline{v}, v_{(1)} \geq \underline{v}] - \mathbb{E}[v_i \mid v_i < \underline{v}] \equiv l > 0.$$

Furthermore, $\max_{j \neq i_1} u_j \geq u_{i_2}$. Then

$$\mathbb{E}\left[\text{ equilibrium revenue } \mid \left(\Pi^{other}, \Pi^{self}\right)\right] - \mathbb{E}[\nu_{(2)}] \leq \left(1 - [q(\lambda)]^N\right)\left(v^K - \mathbb{E}[\nu_{(2)}]\right)$$
$$- [q(\lambda)]^N \Pr\left(\Pi^{other}, \Pi^{self} \mid \sigma_\lambda\right)^N \Pr\left(v_{(1)} \geq \underline{v}, v_{(2)} < \underline{v}\right) \Pr\left(v_{(1)} \geq \underline{v}, v_{(2)} < \underline{v}\right) l.$$

Since $\lim_{\lambda \longrightarrow 0} q(\lambda) = 1$, the first RHS term goes to zero as information costs vanish. Since $\lim_{\lambda \longrightarrow 0} \Pr\left(\Pi^{other}, \Pi^{self} \mid \sigma_\lambda\right) > 0$, the second does not. There exists $\overline{\lambda}$ such that for all $\lambda \leq \overline{\lambda}$, expected revenue is bounded away from the expected second-highest valuation. □

B.3.2 Proof of Proposition 4

Take any equilibrium σ_λ, and fix some realization of $(v_i)_i$. We show that two bidders i_1, i_2 with values $v_{i_2} < v_{i_1}$ cannot both enter the auction with probability greater than $\varepsilon > 0$ for all λ. Hence if two buyers enter with non-vanishing probability, they must have the same value. We then argue that they must be the highest-valuation buyers, i.e., $v_{i_1} = v_{i_2} = \max_j v_j$.

Given an information structure $\left(\Pi^{other}, \Pi^{self}\right)$, let $\pi(v) \in \Pi^{other} \times \Pi^{self}$ denote the information set at which a buyer must be in state $v = (v_i)_i$.[46] Formally, we show that

$$\sum_{\left(\Pi^{other},\Pi^{self}\right)} \Pr\left[\left(\Pi^{other}, \Pi^{self}\right) \mid \sigma_\lambda\right] \Pr\left(\beta(\pi(v)) > 0 \mid \sigma_\lambda\right)$$

cannot be greater than ε irrespective of λ for both i_1 and i_2 if $v_{i_2} < v_{i_1}$. By contradiction, suppose that it is. It must be that at some information sets π_{i_1}, π_{i_2} that have non-vanishing weight given $(v_i)_i$, these two buyers make non-zero bids $\beta(\pi_i) > 0$ for $i = i_1, i_2$.

Step 1. We first show that i_2 must know and bid his valuation at this information set: $\pi_{i_2}^{self} = \{v_{i_2}\}$ and $\beta(\pi_{i_2}) = v_{i_2} + u_{i_2}$. Suppose not: at this information set i_2 does not know his valuation fully $|\pi_{i_2}^{self}| > 1$, and let $\underline{v}$ and $\overline{v}$ denote the lowest and highest values in π_{i_2}, respectively. For i_2 to find it optimal to enter the auction and bid $\beta(\pi_{i_2}) > 0$, it must be that he wins with strictly positive probability at such bid. In particular, this implies that i_1 is sometimes at an information set at which he bids lower than i_2. At information set π_{i_2}, buyer i_2 cannot rule out the possibility that some other buyer $j \neq i_1, i_2$ has the same value as him $v_j = v_{i_2}$. Indeed, $\pi_{i_2}^{other}$ is only informative of $\max_{j \neq i_2} v_j \geq v_{i_1} > v_{i_2}$. Hence with non-vanishing probability, one of i_2's opponents has the same information set as him $\pi_j = \pi_{i_2}$ and makes the same equilibrium bid. With non-vanishing probability, buyer i_2 then ties to win the auction, and must then be indifferent between winning and losing at his equilibrium bid. That equilibrium bid must lie strictly in between $\underline{v} + u_i$ and $\overline{v} + u_i$, which means that there are non-vanishing losses associated with failing to distinguish values $v_i = \underline{v}$ and $v_i = \overline{v}$. Indeed, at the former he does not want to win at his equilibrium bid while he does want to win at the latter. For λ small enough, it must then be optimal to learn to distinguish these values, and it cannot be that $|\pi_{i_2}^{self}| > 1$.

Step 2. We now show that in state $v = (v_i)_i$, agent i_1 must be entering the auction and outbidding i_2 with probability one. Take any information set $\pi(v_{i_1}, v_{-i_1})$ that i_1 might have in equilibrium given the realized $(v_i)_i$. By assumption, one of these information

[46] That is, $\pi(v) = (\{\pi^{other} \in \Pi^{other} : \max_j v_j \in \pi^{other}\}, \{\pi^{self} \in \Pi^{self}(\pi^{other}) : v_i \in \pi^{self}\})$

sets is the one we started with π_{i_1} at which he enters. At this information set, he must be outbidding i_2. Indeed we know from Step 1 that i_2 is bidding $v_{i_2} + u_{i_2} < \nu_{i_1}$, and that i_2 wins sufficiently often at that bid to justify its entry. So there are strictly positive, non-vanishing gains from outbidding i_2, and for λ small enough, i_1 must learn sufficiently about his valuation to do so.

Suppose that at some other information set $\hat{\pi}(v_{i_1}, v_{-i_1})$ that has positive probability in equilibrium, buyer i_1 stays out of the auction in that state of the world, and gets zero gross payoff. By choosing a finer information structure, buyer i_1 could have entered and gotten a strictly positive, non-vanishing gross payoff. Indeed, since i_1 has a strictly greater value than i_2, i_1 must have greater, and hence strictly positive, gains from entering the auction. For λ small enough, that other information structure must yield an overall strictly greater payoff, and hence represents a profitable deviation.

Step 3. We now argue that i_2 cannot find it profitable to enter the auction at information set π_{i_2}. There are two cases: either $\pi_{i_2}^{other}$ is sufficiently fine that i_2 can predict that i_1 will enter with probability one and bid higher than ν_{i_2}, or not. In the first case, i_2 cannot find it optimal to pay the entry fee κ. In the second, i_2 could deviate to a finer $\pi_{i_2}^{other}$ so as to not enter in this state of the world in which there are no gains from doing so. For λ small enough, this strict increase in gross payoff must be lower than the cost of the finer $\pi_{i_2}^{other}$, and this deviation leads to a strictly higher overall payoff.

Taking stocks, given some realization of $(v_i)_i$, if two buyers enter with non-vanishing probability then they must have the same value v_i. We have left to argue that they must have the highest realized value $\max_i v_i$. By contradiction, suppose not: two buyers i_1 and i_2 enter with non-trivial probability while some other buyers j^* with $v_{j^*} > v_{i_1} = v_{i_2}$ does not. First note that bidders i_1 and i_2 must know their valuation fully for λ small enough and bid it, as they tie against each other with non-vanishing probability. Yet if they find it profitable to enter and bid their values then buyer j^* must find it strictly profitable since he has a strictly higher valuation than them. Hence j^* cannot stay away from the auction in equilibrium, and if two buyers enter with non-vanishing probability then they must have $v_i = \max_j v_j$. □

B.4 Proof of Results of Section 5

B.4.1 Proof of Proposition 5

Let $\Pi^{other} = \{\{\pi_l^{other}\}_{l=1}^{L}\}$ be an information partition about $\max_j v_j$ that has non-vanishing probability in equilibrium. Let $v^{\underline{k}_l}$ and $v^{\bar{k}_l}$ denote the lowest and highest values in π_l^{other}, respectively. The equilibrium partition may depend on the number of bidders N but we do not make that dependence explicit so as to keep the notation uncluttered. The proof of Theorem 5 has two steps. We first show that, for N large enough and λ small enough, buyers learn to distinguish whether or not their toughest competitor has a value equal to v^K (i.e., $\{v^K\} \in \Pi^{other}$) under any information structure that has non-vanishing weight. Second, we show that setting a reserve price just below that highest possible valuation yields more revenue than having an additional buyer participate in the auction.

Step 1. *Assume no reserve price. There exists $\overline{N}_1$ such that, for all $N \geq \overline{N}_1$, there exists $\bar{\lambda}$ such that, for all $\lambda \leq \bar{\lambda}$, buyers learn whether their toughest competitor has the highest possible value: $\{v^K\} \in \Pi^{other}$.*

By contradiction, suppose not: for arbitrarily large N, there exists an information structure that has non-vanishing weight in equilibrium such that $\{v^K\} \notin \Pi^{other}$. That is, with non-trivial probability, buyers choose a partition about others that bundles v^K with some other possible values: $\{v^{\underline{k}_L}, \ldots, v^K\} \in \Pi^{other}$ for some $v^{\underline{k}_L} < v^K$.

We know from Theorem 1 that an information structure that has non-vanishing weight in equilibrium must satisfy ($\star$) and minimize cost. We give an alternative information structure under which $\{v^K\} \in \widehat{\Pi}^{other}$ that is strictly cheaper than $\left(\Pi^{other}, \Pi^{self}\right)$, which hence proves the latter cannot be sustained in equilibrium.

The cost of information partition Π^{other} equals:

$$c\left(\Pi^{other}, p_{1:N-1}\right) = H\left(p_{1:N-1}\right) - \sum_{l=1}^{L} \Pr\left(\max_j v_j \in \pi_l^{other}\right) H\left(\delta_{\pi_l^{other}}\right),$$

where $\delta_{\pi_l^{other}} \in \Delta V$ is the agent's posterior about $\max_j v_j$ given that he knows $\max_j v_j \in \pi_l^{other}$:

$$\delta_{\pi_l^{other}}(\cdot) = \Pr\left(\cdot \,\middle|\, \max_j v_j \in \pi_l^{other}\right) = \begin{cases} \frac{p_{1:N-1}(v)}{\sum_{v' \in \pi_l^{other}} p_{1:N-1}(v')} & \text{if } v \in \pi_l^{other} \\ 0 \quad \text{otherwise} \end{cases}.$$

Consider an alternative information partition about others that unbundles v^K from $\{v^{\underline{k}_L}, \dots, v^{K-1}\}$, but keeps the rest of the partition the same:

$$\widehat{\Pi}^{other} = \{\{v^{\underline{k}_l}, \dots, v^{\overline{k}_l}\}_{l=1,\dots,L-1}, \{v^{\underline{k}_L}, \dots, v^{K-1}\}, \{v^K\}\}.$$

Naturally this partition is finer, and hence costlier than the one we started with. Using the same formula as above, the extra cost equals

$$c(\widehat{\Pi}^{other}, p_{1:N-1}) - c(\Pi^{other}_\lambda, p_{1:N-1}) = \Pr\left(\max_j v_j \geq v^{\underline{k}_L}\right) H\left(\delta_{\{v^{\underline{k}_L},\dots,v^K\}}\right) \\ - \Pr\left(v^{\underline{k}_L} \leq \max_j v_j < v^K\right) H\left(\delta_{\{v^{\underline{k}_L},\dots,v^{K-1}\}}\right) - \Pr\left(\max_j v_j = v^K\right) H\left(\delta_{\{v^K\}}\right).$$

Note that as N goes to infinity, the second term goes to zero, as $\Pr\left(\max_j v_j < v^K\right)$ goes to zero. Furthermore, a buyer's posterior conditional on learning that $\max_j v_j \in \{v^{\underline{k}_L}, \dots, v^K\}$ converges to the degenerate belief that puts probability one on the toughest competitor having the highest possible value $\delta_{\{v^K\}}$. Indeed, as N goes to infinity, the probability of such event goes to one. Hence, $\lim_{N\longrightarrow\infty} \delta_{\{v^{\underline{k}_L},\dots,v^K\}} = \delta_{\{v^K\}}$. Since H is continuous, this implies that the first and last term cancel out in the limit, and so the overall expression goes to zero: the cost of learning whether one of the competitors has the highest valuation possible becomes negligible.[47]

There is however a first-order gain in choosing this alternative partition, as it allows the buyer to save on information cost about his own valuation. If the buyer learns that $\max_j v_j \in \{v^{\underline{k}_L}, \dots, v^K\} = \pi^{other}$, then we know from Theorem 1 that $\Pi^{self}(\pi^{other}) = \{\{v^1, \dots, v^{\overline{k}_{L-1}}\}, \{v^k\}_{k=\underline{k}_L,\dots,K}\}$. If the buyer learns that $\max_j v_j \in \{v^{\underline{k}_L}, \dots, v^{K-1}\}$, then the optimal way to partition his set of valuations is the same one. If however the buyer learns that $\max_j v_j = v^K$, then he optimally chooses a coarser partition for himself: $\widehat{\Pi}^{self}(\{v^K\}) = \{\{v^1, \dots, v^{K-1}\}, \{v^K\}\}$. Hence the gain in information cost on self is

$$\Pr\left(\max_j v_j = v^K\right) \\ \times \left[c\left(\{\{v^1, \dots, v^{K-1}\}, \{v^K\}\}, p\right) - c\left(\{\{v^1, \dots, v^{\tau_L}\}, \{v^k\}_{k=\tau_L+1,\dots,K}\}, p\right)\right].$$

Since $\Pr(\max_j v_j = v^K)$ tends to one as N tends to infinity, this tends to the strictly

[47] Intuitively, this is due to the fact that such event is so likely that learning about its realization does not move the buyer's belief much.

positive expression in parenthesis. For N large enough, the information structure $(\widehat{\Pi}^{other}, \widehat{\Pi}^{self})$ is strictly cheaper, and so the information structure we started with cannot have non-trivial weight in equilibrium.

Step 2. *There exists $\overline{N}_2$ such that, for all $N \geq \overline{N}_2$, there exists $\bar{\lambda}$ such that, for all $\lambda \leq \bar{\lambda}$, setting a reserve price $r \in (v^{K-1}, v^K)$ yields more revenue than having one more bidder in the auction.*

Consider a (somewhat extreme) reserve price that lies just below the highest possible valuation $r = v^K - \eta$ for η small ($|\underline{u}_N| < \eta < v^K - v^{K-1}$). Under such reserve price, and for λ small enough, all buyers find it optimal to acquire no information about others, and to only learn whether their valuations lie above or below the reserve price. Hence they all choose

$$\Pi^{self} = \{\{v^1, v^2, \ldots, v^{K-1}\}, \{v^K\}\}.$$

Given a realization of $(v_i)_i$, let $v_{1:N}$ and $v_{2:N}$ be the highest and second-highest valuations, respectively. As λ goes to zero, imposing a reserve price $r = v^K - \eta$ then yields an expected revenue of

$$\text{Revenue}(r, N) = \Pr\left(v_{1:N} = v^K, v_{2:N} < v^K\right) r + \Pr\left(v_{2:N} = v^K\right) \mathbb{E}\left(\nu_{2:N} \mid \nu_{2:N} \geq v^K + \underline{u}_N\right).$$

Now consider what happens in equilibrium if no reserve price is imposed, but there are $N+1$ bidders participating in the auction. By Step 1, it has to be that buyers choose to learn sufficiently finely about the competition, and in particular that, for N large enough, they come to learn whether their toughest competitor has a value of v^K. If they learn that this is the case, they then partition their own set of valuations into $\{\{v^1, v^2, \ldots, v^{K-1}\}, \{v^K\}\}$. That is, they learn whether they should compete with their toughest competitor (which only yields a non zero payoff when they also have $v_i = v^K$) and bundle together all valuations below v^K. Hence, in any equilibrium, all buyers who learn $\max_j v_j = v^K$ and $v_i < v^K$ bid $\mathbb{E}[v_i | v_i < v^K] + u_i$ (as argued in Step 1 of proof of Theorem 2).

Revenue with $N+1$ bidders but no reserve price then equals

$$\begin{aligned}\text{Revenue}(0, N+1) = {} & \Pr\left(v_{2:N+1} = v^K\right) \mathbb{E}\left(\nu_{2:N+1} \mid \nu_{2:N+1} \geq v^K + \underline{u}_{N+1}\right) \\ & + \Pr\left(v_{1:N+1} = v^K, v_{2:N+1} < v^K\right)\left(\mathbb{E}\left[v_i | v_i < v^K\right] + \mathbb{E}(u_{1:N})\right)\end{aligned}$$

$$+ \Pr\left(v_{1:N+1} < v^K\right) \mathbb{E}\left[\text{eq. revenue}|v_i < v^K\, \forall i = 1, \ldots, N+1\right].$$

The first line captures expected revenue when both the highest- and second-highest-valuation buyers have $v_i = v^K$. In such case they both learn their valuations fully, and revenue simply equals the expected second-highest valuation. The second line captures expected revenue when the highest-valuation buyer has $v_i = v^K$ while the second-highest valuation buyer has $v_j < v^K$. In such case, all losing buyers j fail to learn their valuations and bid $\mathbb{E}\left[v_i|v_i < v^K\right] + u_j$. Finally, the third line captures expected revenue when none of the buyers has $v_i = v^K$. The probability of such event is vanishing at a faster rate than the others as N grows. We will thus be able to overlook it and do not need to derive an explicit expression for revenue.

We show that, for N high enough, the above reserve price yields greater expected revenue than having an additional bidder:

$$\Delta \equiv \text{Revenue}(r, N) - \text{Revenue}(0, N+1) > 0.$$

This difference is at least

$$\begin{aligned}
\Delta \geq\ & \Pr(v_{1:N} = v^K, v_{2:N} < v^K) \left(r - \mathbb{E}[v_i|v_i < v^K] - \overline{u}_{N+1}\right) \\
& + \left(\Pr(v_{2:N} = v^K) - \Pr(v_{2:N+1} = v^K)\right) (v^K + \overline{u}_N) \\
& + \left(\Pr(v_{1:N} = v^K, v_{2:N} < v^K) - \Pr(v_{1:N+1} = v^K, v_{2:N+1} < v^K)\right) \left(\mathbb{E}[v_i|v_i < v^K] + \overline{u}_{N+1}\right) \\
& - \Pr(v_{1:N+1} < v^K)\mathbb{E}[\text{eq. revenue}|v_i < v^K\, \forall i = 1, \ldots, N+1] \\
& + \Pr\left(v_{2:N} = v^K\right) \Big[\mathbb{E}\left(u_{2:M}|M = |\{i = 1, \ldots, N|v_i = v^K\}|, M \geq 2\right) \\
& \qquad\qquad - \mathbb{E}\left(u_{2:M}|M = |\{i = 1, \ldots, N+1|v_i = v^K\}|, M \geq 2\right)\Big] \\
=\ & Np^K \left(1 - p^K\right)^{N-1} \left(r - \mathbb{E}[v_i|v_i < v^K] - \overline{u}_{N+1}\right) \\
& + (v^K + \overline{u}_N)\Big[1 - \left(1 - p^K\right)^N - Np^K \left(1 - p^K\right)^{N-1} - 1 + \left(1 - p^K\right)^{N+1} + (N+1)p^K \left(1 - p^K\right)^N\Big] \\
& + \Big[Np^K \left(1 - p^K\right)^{N-1} - (N+1)p^K \left(1 - p^K\right)^N\Big] \left(\mathbb{E}[v_i|v_i < v^K] + \overline{u}_{N+1}\right) \\
& - \left(1 - p^K\right)^{N+1} \mathbb{E}[\text{eq. revenue}|v_i < v^K\, \forall i = 1, \ldots, N+1] \\
& + \Pr\left(v_{2:N} = v^K\right) \Delta_u^N,
\end{aligned}$$

where

$$\Delta_u^N = \mathbb{E}\left(u_{2:M} | M = |\{i = 1, \ldots, N | v_i = v^K\}|, M \geq 2\right)$$
$$- \mathbb{E}\left(u_{2:M} | M = |\{i = 1, \ldots, N+1 | v_i = v^K\}|, M \geq 2\right).$$

Factorizing by $Np^K \left(1 - p^K\right)^{N-1}$ this simplifies to

$$\begin{aligned}\frac{\Delta}{Np^K \left(1 - p^K\right)^{N-1}} \geq r - \mathbb{E}[v_i | v_i < v^K] - \overline{u}_{N+1} - p^K(v^K + \overline{u}_N) \\ + \left[p^K - \frac{1 - p^K}{N}\right] \left[\mathbb{E}[v_i | v_i < v^K] + \overline{u}_{N+1}\right] \\ - \frac{\left(1 - p^K\right)^2}{Np^K} \mathbb{E}[\text{eq. revenue} | v_i < v^K \, \forall i = 1, \ldots, N+1] \\ + \Pr\left(v_{2:N} = v^K\right) \frac{\Delta_u^N}{Np^K \left(1 - p^K\right)^{N-1}}.\end{aligned}$$

Using the fact that $r = v^K - \eta$ for some small η, this rewrites as

$$\begin{aligned}\frac{\Delta}{Np^K \left(1 - p^K\right)^{N-1}} \geq \left(1 - p^K\right) \left(v^K - \mathbb{E}[v_i | v_i < v^K]\right) - \eta - p_K \overline{u}_N - (1 - p_K) \overline{u}_{N+1} \\ - \frac{1 - p^K}{N} \left(\mathbb{E}[v_i | v_i < v^K] + \overline{u}_{N+1}\right) - \frac{\left(1 - p^K\right)^2}{Np^K} \mathbb{E}[\text{eq. revenue} | v_i < v^K \, \forall i] \\ + \Pr\left(v_{2:N} = v^K\right) \frac{\Delta_u^N}{Np^K \left(1 - p^K\right)^{N-1}}.\end{aligned}$$

First note that $(1-p^K) \left(v^K - \mathbb{E}[v_i | v_i < v^K]\right) > 0$ and so must be greater than η for η small enough. Furthermore, all terms on the second line go to zero as N goes to infinity. The same is true of the bounds on the noise terms $\overline{u}_N$. Finally, the last term is bounded above by $\overline{u}_N - \underline{u}_N$ and is arbitrarily small if the noise terms have small support. Hence, there must exist $\bar{N}_2$ such that, for all $N \geq \bar{N}_2$, the reserve price yields greater revenue than the extra bidder: $\Delta > 0$.

Intuitively, for large N, the first-order difference between a reserve price and an additional bidder occurs when only one bidder has value $v_i = v^K$ and all others have values $v_i < v^K$.[48] In such a case, an additional bidder only yields revenue $\mathbb{E}[v_i | v_i < v^K] + \max_j u_j$ whereas a reserve price yields revenue arbitrarily close to $v_K + \underline{u}_N$, which is higher.

Combining Steps 1 and 2, for all $N \geq \max\{\bar{N}_1, \bar{N}_2\}$ the claim holds, and the above reserve price outperforms an auction with no reserve but one more bidder. $\square$

B.4.2 Proof of Proposition 6 and Proposition 7

Proof of Proposition 6. Consider the following way to randomize access to the auction:

$$\Pr(M = N) = 1 - q \quad \text{and} \quad \Pr(M = N \setminus i) = \frac{q}{N} \text{ for all } i,$$

for some $q \in (0, 1)$. That is, with probability $1 - q$ all buyers get access to the auction. With remaining probability, one buyer chosen uniformly at random is excluded. Take any information partition about self Π^{self} that has non-vanishing weight in equilibrium. We show that, for any $q > 0$, there exists $\bar{\lambda}$ such that, for all $\lambda \leq \bar{\lambda}$, buyers must always become fully informed of their valuations: $\Pi^{self} = \{\{v\}_{v \in V}\}$.

By contradiction, suppose not: Π^{self} bundles some valuations together. Take any such bundle and let $\underline{v}$ and $\overline{v}$ be the lowest and highest element in that bundle, respectively. Denote that bundle by $\pi^{self}_{\underline{v},\overline{v}}$ and let $\beta(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i)$ a buyer's equilibrium bid at that bundle. We know that $\beta(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i) \in [\underline{v} + u_i, \overline{v} + u_i]$.

Consider deviating to $\widehat{\Pi}^{self} = \{\{v\}_{v \in V}\}$. Since that partition is finer, it is costlier and increases buyer i's information costs by

$$\lambda \left[c\left(\widehat{\Pi}^{self}, p \right) - c\left(\Pi^{self}, p \right) \right] > 0.$$

Such a finer partition must (weakly) increase buyer i's gross payoff. We show that it strictly increases i's gross payoff. Whatever information i acquired about his toughest opponent's valuation is irrelevant with probability q/N. Indeed, with probability q/N, i's toughest opponent will be excluded from the auction. In such event, buyer i is left competing with buyers about whose valuations i has no information. In particular, with non-trivial probability, all these buyers chose the same signal about self Π^{self}, all

[48] When more than one bidders have values $v_i = v^K$, they learn their valuations fully and the second-highest bid is $\nu_{2:N} > r$. The only difference in revenue between a reserve price and an additional bidder then comes from the fact that, under the latter, the second-highest bidder has a larger draw of u_i in expectation.

have values $v_j \in \pi^{self}_{\underline{v},\overline{v}}$, and all bid $\beta(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_j)$ for some u_j. But buyer i then has a strict incentive to learn to distinguish $v_i = \overline{v}$ from $v_i = \underline{v}$. Indeed, he wants to win against $\beta(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, \max_j u_j)$ (at least) when $v_i = \overline{v}$ and $u_i > \max_j u_j$, and lose against $\beta(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, \max_j u_j)$ (at least) when $v_i = \underline{v}$ and $u_i < \max_j u_j$.

These gains from distinguishing $v_i = \overline{v}$ and $v_i = \underline{v}$ are strictly positive and independent of λ. Hence, for λ small enough, the increase in information cost must be strictly smaller than the gains, and buyers must become fully informed about their valuations in equilibrium: $\Pi^{self} = \{\{v\}_{v \in V}\}$.

In any equilibrium, fully informed buyers must bid their valuations for the good: $\beta((\pi^{other}, \{v_i\}, u_i)) = v_i + u_i = \nu_i$. Expected revenue is then at least $q\mathbb{E}[\nu_{(2)}]$. Given any $\varepsilon > 0$, set $q = 1 - \frac{\varepsilon}{\mathbb{E}[\nu_{(2)}]}$. For λ small enough, expected revenue from such randomized access is at least $\mathbb{E}[\nu_{(2)}] - \varepsilon$, which completes the proof. □

Proof of Proposition 7. Consider the randomized access from Proposition 6. For any $\varepsilon > 0$, the seller can ensure himself a revenue of

$$\mathbb{E}[\nu_{(2)}] - \varepsilon$$

for λ small enough. We compare this revenue to the one obtained when the seller sets the optimal reserve price.

Step 1. We first show that, for N large enough and λ small enough, the optimal reserve price r^* must target buyers with the highest possible realization of v_i, that is, $r^* = v^K + u^*$ for some $u^* \in (\underline{u}_N, \overline{u}_N)$. Suppose not, and $r^* < v^K + \underline{u}_N$ for all N and λ. Using the same argument as in the proof of Theorem 5, we show that, for N large enough, buyers only put non-vanishing weight on information structures such that $\{v^K\} \in \Pi^{other}$. In words, they find it cost-efficient to learn whether they face a competitor with the highest possible valuation, as this event has very high probability for large N. (See proof of Theorem 5 for more details.)

After learning that $\max_j v_j = v^K$, a buyer chooses $\Pi^{self}(\{v^K\}) = \{\{v_i : v_i < v^K\}, \{v^K\}\}$ and bids $\mathbb{E}[v_i \mid v_i < v^K] + u_i$ whenever $v_i < v^K$. A reserve price of $r < v^K + \underline{u}_N$ then yields revenue strictly bounded above by

$$\left(1 - \left(1 - p^K\right)^N - Np^K\left(1 - p^K\right)^{N-1}\right)\mathbb{E}[\nu_{(2)} \mid \nu_{(2)} \geq v^K + \underline{u}_N]$$

$$+ Np^K \left(1-p^K\right)^{N-1} \mathbb{E}\left[\max\{r, \mathbb{E}[v_i \mid v_i < v^K] + u_i\}\right] + \left(1-p^K\right)^N \left(v^K + \underline{u}_N\right).$$

In comparison, a reserve price to $r' = v^K + \underline{u}_N$ yields an expected revenue of

$$\left(1 - \left(1-p^K\right)^N - Np^K \left(1-p^K\right)^{N-1}\right) \mathbb{E}[\nu_{(2)} \mid \nu_{(2)} \geq v^K + \underline{u}_N]$$
$$+ Np^K \left(1-p^K\right)^{N-1} \left(v^K + \underline{u}_N\right).$$

The latter is strictly greater whenever

$$Np^K \left(1-p^K\right)^{N-1} \left(v^K + \underline{u}_N - \mathbb{E}\left[\max\{r, \mathbb{E}[v_i \mid v_i < v^K] + u_i\}\right]\right)$$
$$- \left(1-p^K\right)^N \left(v^K + \underline{u}_N\right) > 0$$
$$\iff Np^K \left(v^K + \underline{u}_N - \mathbb{E}\left[\max\{r, \mathbb{E}[v_i \mid v_i < v^K] + u_i\}\right]\right) - \left(1-p^K\right) \left(v^K + \underline{u}_N\right) > 0.$$

Since $\left[\max\{r, \mathbb{E}[v_i \mid v_i < v^K] + u_i\}\right] \leq r < v^K + \underline{u}_N$, the first term is strictly positive, and for N large enough the inequality must hold. The optimal reserve price must then lie weakly above $v^K + \underline{u}_N$.

Step 2. We now show that revenue under the optimal reserve price is lower than under the above randomized access. Under the optimal reserve price, expected revenue equals

$$r^* N \left(p^K \Pr(u_i \geq u^*)\right) \left(1 - p^K \Pr(u_i \geq u^*)\right)^{N-1}$$
$$+ \left[1 - \left(1 - p^K \Pr(u_i \geq u^*)\right)^N - N \left(p^K \Pr(u_i \geq u^*)\right) \left(1 - p^K \Pr(u_i \geq u^*)\right)^{N-1}\right]$$
$$\times \mathbb{E}[\nu_{(2)} \mid \nu_{(2)} \geq r^*].$$

The first term arises whenever the reserve price binds, which is the case whenever only one buyer has a valuation greater than $r^* = v^K + u^*$. The second term captures expected revenue when the reserve price does not bind, that is when two or more buyers have valuations $\nu_i = v_i + u_i \geq v^K + u^*$.

Randomizing access yields revenue $\mathbb{E}[\nu_{(2)}] - \varepsilon$. This is greater whenever

$$\left[\left(1 - p^K \Pr(u_i \geq u^*)\right)^N + N \left(p^K \Pr(u_i \geq u^*)\right) \left(1 - p^K \Pr(u_i \geq u^*)\right)^{N-1}\right] \mathbb{E}[\nu_{(2)} \mid \nu_{(2)} \leq r^*]$$
$$- r^* N \left(p^K \Pr(u_i \geq u^*)\right) \left(1 - p^K \Pr(u_i \geq u^*)\right)^{N-1} - \varepsilon > 0.$$

This rewrites as

$$\left(1-p^K \Pr(u_i \geq u^*)\right)^N \mathbb{E}\left[\nu_{(2)} \mid \nu_{(1)} \leq r^*\right] - \varepsilon$$
$$- N\left(p^K \Pr(u_i \geq u^*)\right)\left(1-p^K \Pr(u_i \geq u^*)\right)^{N-1}\left(r^* - \mathbb{E}[\nu_{(2)} \mid \nu_{(2)} \leq r^*, \nu_{(1)} > r^*]\right) > 0.$$

In words, the gain from randomizing access comes from the fact that buyers fully learn their valuations, and that whenever none of them has a value above the reserve price, then the seller gets a revenue equal to the second-highest value (first line). The loss comes from states of the world in which only one buyer has a value above the reserve price (so only one buyer has $\nu_i \geq v^K + u^*$), in which case the reserve price yields a revenue of $r^* = v^K + u^*$ while the randomization into entry yields a revenue of $\mathbb{E}[\nu_{2:N} \mid \nu_{2:N} \leq v^K + u^*, \nu_{1:N} > v^K + u^*] - \varepsilon = \mathbb{E}[\nu_{1:N-1} \mid \nu_{1:N-1} \leq v^K + u^*] - \varepsilon$ (second line). However this loss goes to zero with N at a faster rate than the gains, as the expected highest-value below the reserve price converges to the reserve price quickly. In particular, with probability $1-[(1-p^K)/(1-p^K \Pr(u_i \geq u^*))]^{N-1}$, the second-highest value is $\nu_{2:n} = v^K + u_i$ for some $u_i < u^*$. Hence the condition rewrites as

$$\left(1-p^K \Pr(u_i \geq u^*)\right)^N \mathbb{E}\left[\nu_{(2)} \mid \nu_{(1)} \leq r^*\right] - \varepsilon$$
$$- N\left(p^K \Pr(u_i \geq u^*)\right)\left(1-p^K \Pr(u_i \geq u^*)\right)^{N-1}\left[1-\left(\frac{1-p^K}{1-p^K \Pr(u_i \geq u^*)}\right)^{N-1}\right]$$
$$\times\left(u^* - \mathbb{E}[u_{1:M} \mid M = |\{i : v_i = v^K, u_i < u^*\}|]\right)$$
$$- N\left(p^K \Pr(u_i \geq u^*)\right)\left(1-p^K \Pr(u_i \geq u^*)\right)^{N-1}\left(\frac{1-p^K}{1-p^K \Pr(u_i \geq u^*)}\right)^{N-1}$$
$$\times\left(v^K + u^* - \mathbb{E}[\nu_{1:N-1} \mid \nu_{1:N-1} \leq v^K + \underline{u}_N]\right) > 0.$$

Dividing everything by $\left(1-p^K \Pr(u_i \geq u^*)\right)^{N-1}$ yields

$$\left(1-p^K \Pr(u_i \geq u^*)\right) \mathbb{E}\left[\nu_{(2)} \mid \nu_{(1)} \leq r^*\right] - \varepsilon$$
$$- N\left(p^K \Pr(u_i \geq u^*)\right)\left[1-\left(\frac{1-p^K}{1-p^K \Pr(u_i \geq u^*)}\right)^{N-1}\right]$$
$$\times\left(u^* - \mathbb{E}[u_{1:M} \mid M = |\{i : v_i = v^K, u_i < u^*\}|]\right)$$

$$
- N\left(p^K \Pr(u_i \geq u^*)\right)\left(\frac{1-p^K}{1-p^K \Pr(u_i \geq u^*)}\right)^{N-1} \\
\times \left(v^K + u^* - \mathbb{E}[\nu_{1:N-1} \mid \nu_{1:N-1} \leq v^K + \underline{u}_N]\right) > 0.
$$

For ε small enough, the first term is strictly positive and remains bounded away from zero as N grows large. The last term goes to zero as N grows large. So does the second term if the noise terms are sufficiently small given N since it scales with $u^* - \underline{u}_N \leq \overline{u}_N - \underline{u}_N$. Hence there exists $\bar{N}$ such that, for all $N \geq \bar{N}$, the claim holds: the above randomization intro entry outperforms an auction where all bids are considered but with an optimal reserve price. $\square$

B.5 Proof of Results in Section 6

Proof of Proposition 8. We show that there cannot exist an equilibrium in which buyers first choose to learn about their own values and then about others.

By contradiction, suppose such an equilibrium exists, and let Π^{self} be an information partition about self that has non-vanishing probability in equilibrium. We first argue that, for sufficiently small information costs λ, $\Pi^{self} = \{\{v\}_{v \in V}\}$. Suppose not, such that there exists $\pi^{self} \in \Pi^{self}$ with $|\pi^{self}| > 1$. Let $\underline{v}$ and $\overline{v}$ be the smallest and highest values in π^{self}, respectively. In states of the world where $v_j \in \pi^{self}$ for all buyers j, all buyers are at the same information set π^{self} with non-vanishing probability. Following a similar argument as in the proof of Lemma 3, they then tie for the good with non-vanishing probability, and must be indifferent between losing and winning at their equilibrium bid. Their equilibrium bid then lies strictly in between $\underline{v} + u_i$ and $\overline{v} + u_i$, and given that they face such bid with non-vanishing probability, they have a strict incentive to learn to distinguish value $v_i = \underline{v}$ from value $v_i = \overline{v}$. Therefore a buyer learns his value fully $\Pi^{self} = \{\{v_i\}_{v_i \in V}\}$. We however know from Proposition 3 that such an equilibrium cannot exist. $\square$

Proof of Proposition 9. The proof of Proposition 3 is identical. Indeed, it proceeds by contradiction: we assume that an equilibrium in which buyers converge to fully learning their values exists and then show that they have a strict incentive to deviate to an alternative information structure. That deviation still exists in this extended model since we are only expanding the set of information structures available to buyers, and so the conclusion follows directly.

The proof of Theorem 3 extends almost verbatim as well. During the information acquisition stage, buyers can acquire any sequence of partitions about their values and others'. Irrespective of the exact form this process takes, their information set upon entering the auction is fully described by an interval of values they deem possible for themselves π_i^{self} and an interval of values they deem possible for their toughest opponent π_i^{other}. Take any information structure that had non-vanishing weight in equilibrium. The same arguments as in the proof of Lemma 3 yield that if these two intervals overlap, then the belief about self must be degenerate ($|\pi_i^{self}| = 1$). This ensures buyers make no non-vanishing mistakes in equilibrium as they learn their valuations fully when the auction is "close." Furthermore, the same arguments as in the proof of Lemma 6 yield that buyers must choose the cheapest such information structure. We know that fully learning one's value if not cost-efficient, so buyers must acquire some information partition about others at some point in the learning process. In particular, this means that if at any stage in the information acquisition process, a buyer knows for sure that his value is above (or below) that of his toughest opponent, then it cannot be cost-efficient for that buyer to learn more. So any non-vanishing information structure must have losing buyers fail to learn their valuation with positive probability, and the proof of Theorem 2 applies. □

Proof of Theorem 3. The following two facts simplify the equilibrium analysis. First, buyers must bid their expected valuation given their information set in any equilibrium. Indeed, doing so is a weakly dominant strategy: others' bids no longer carry any relevant information to them and so the standard strategy-proofness of the second price auction applies. Furthermore, bidding truthfully is the only bidding strategy that survives our tremble refinement. Second, it is without loss to look at pure learning strategies. Indeed, any mixing over distribution over posterior beliefs can be replicated by a single distribution over posterior beliefs that has the same cost. Taken together, these two facts imply that any equilibrium can be fully described by the chosen signal about others $\Pi^{other} : \{\underline{s}, \overline{s}\} \longrightarrow \Delta\Delta(\{\underline{s}, \overline{s}\})$ and about self $\Pi^{self} : [0, 1] \times \{\underline{s}, \overline{s}\} \longrightarrow \Delta\Delta V$.[49]

Take any sequence of equilibria $\{(\Pi_\lambda^{other}, \Pi_\lambda^{self})\}_\lambda$ and denote the limiting information structure by $(\Pi^{other}, \Pi^{self}) \equiv \lim_{\lambda \longrightarrow 0}(\Pi_\lambda^{other}, \Pi_\lambda^{self})$.

Step 1: We show that $\operatorname{supp} \Pi^{other}(\underline{s}) = \{\underline{\mu}, 1\}$ and $\operatorname{supp} \Pi^{other}(\overline{s}) = \{0, \overline{\mu}\}$. In words,

[49]Recall that buyers get to observe s_i for free at the beginning of the game so they can condition their learning strategy on the realization of that signal.

a buyer who knows his value is low ($s_i = \underline{s}$) will choose a signal about others that either reveals the competition is strong for sure (posterior probability that $\max_j s_j = \overline{s}$ is 1) or not. Similarly, a buyer who knows his value is high ($s_i = \overline{s}$) will choose a signal about others that either reveals the competition is weak for sure (posterior probability that $\max_j s_j = \overline{s}$ is 0) or not. To that end, we first argue that the good must be allocated to the highest valuation bidder under the limiting information structure $(\Pi^{other}, \Pi^{self})$. Indeed, if the equilibrium allocation were not efficient, then some buyer would have a strict incentive to deviate: since buyers bid their expected value given their information set, any misallocation must be driven by some buyer having failed to learn his valuation fully. But then that buyer would be leaving money on the table, as he would either win at a price above his value or lose against a price below his value. In the limit, learning his value fully and avoiding this mistake must be strictly profitable.

For the good to be efficiently allocated in the limit, it must be that $\max_j s_j = s_i$ $\Pi^{self}(\overline{s})[\mu] = \{\delta_{\tilde{v}}\}_{\tilde{v}>v^*}$ whenever $\mu > 0$ and $\Pi^{self}(\underline{s})[\mu] = \{\delta_{\tilde{v}}\}_{\tilde{v}\leq v^*}$ whenever $\mu < 1$, where $\{\delta_{\tilde{v}}\}_{\tilde{v}}$ denotes the "corner" distribution over posterior beliefs under which a buyer fully learns his valuation. In words, for the good to be allocated, it must be that a buyer with $s_i = \overline{s}$ fully learns his value whenever he assigns probability to the event $\max_j s_j = \overline{s}$. Similarly, it must be that a buyer with $s_i = \underline{s}$ fully learns his value whenever he assigns probability to the event $\max_j s_j = \underline{s}$. A direct implication is that, if $1 \notin \operatorname{supp}\Pi^{other}(\underline{s})$, then a buyer with $s_i = \underline{s}$ would *always* learn his valuation fully at the second stage of the information acquisition process. But then learning anything about the competition would be pointless and buyer i would simply learn nothing about the competition and everything about his own value. However, if his prior is sufficiently uncertain this would not be cost-efficient, and he would strictly reduce his overall information cost by learning fully the realization of $\max_j s_j$ and his value only if $\max_j s_j = s_i$. (The argument for this is the same as in the proof of Proposition 3.) Thus $1 \in \operatorname{supp}\Pi^{other}(\underline{s})$. Finally, it cannot be that $\mu', \mu'' \in \operatorname{supp}\Pi^{other}$ with $\mu' < 1$ and $\mu'' < 1$. Indeed, at both of these posterior buyer i would fully learn his valuation. Merging these two posteriors would strictly reduce his information costs without changing his bidding strategy, and so doing it is a profitable deviation. Overall that means $\operatorname{supp}\Pi^{other}(\underline{s}) = \{\underline{\mu}, 1\}$ for some $\mu < 1$. The proof of $\operatorname{supp}\Pi^{other}(\overline{s}) = \{0, \overline{\mu}\}$ is virtually identical.

Step 2: We show that expected revenue under $(\Pi^{other}, \Pi^{self})$ is strictly below the expected second-highest valuation. Recall that whenever buyers learn their valuations fully, they bid their valuations, and revenue equals to second-highest value. Hence realized revenue can only differ from the realized second-highest value when some losing buyer(s) failed to learn their valuation(s) fully. This can only happen when the second-highest bid is set by a buyer i with $s_i = \underline{s}$ who learned that his toughest opponent has $\max_j s_j = \overline{s}$ and, as a result, chose to learn nothing about his value. The second-highest bid is then $\mathbb{E}[v \mid v \leq v^*] + u_i$. Following the same argument as in the proof of Theorem 2, that second-highest bid is strictly lower than if i had learned his value fully and bid his value as long as $N > 2$. $\square$