How Competition Shapes Information in Auctions*

Agathe PERNOUD† Simon GLEYZE‡

December 30

Click here for the latest version.

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. It also undermines price discovery. Finally, we examine the implications for auction design. First, setting an optimal reserve price is more valuable than attracting an extra buyer, which contrasts with the seminal result of Bulow and Klemperer (1996). Second, the seller can incentivize buyers to learn their valuations, hence restoring effective competition, by maintaining uncertainty over the set of auction participants.

*We are grateful to Matthew Jackson and Paul Milgrom for invaluable guidance and support. We also thank Mohammad Akbarpour, Piotr Dworczak, Liran Einav, Ellen Muir, Doron Ravid, Ilya Segal, Andy Skrzypacz, and seminar participants at Stanford for helpful discussions, comments, and questions. 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.

†Department of Economics, Stanford University. Email: agathep@stanford.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 submitting a bid. Relevant examples are the sales of complex, high-value assets such as companies, broadband licenses, or procurement contracts, during which interested buyers conduct extensive due diligence. For instance, bidders in takeover auctions get access to extensive information about the target company's operations and finances, allowing them to assess any synergies and estimate how much they value its acquisition.

In practice, accessing and processing such information is costly, and buyers only want to undertake this investment if they have a fair chance of winning the auction.[1] They then have an incentive to investigate how much other participants value the good for sale: doing so reduces the strategic uncertainty faced in the auction and allows them to assess whether due diligence costs are worth paying. In Gleyze and Pernoud (2022), we show that such an incentive is prevalent and arises under most auction formats. We now argue that it significantly affects the properties of auctions and address the following questions: How does competition between buyers shape the kinds of information they acquire? How does that, in turn, affect the value of competition in auctions?

Answering these questions is key to understanding the performance of some real-life auctions. For example, consider the 1994-1995 spectrum auction run by the Federal Communications Commission. The ascending auction resulted in a low price of $26 per capita for the Los Angeles license, while less profitable licenses were sold at higher prices.[2] Some argue that the presence of Pacific Telesis, a company that already operated in California and was presumed to win, scared away its competitors. Indeed, other participants came to learn PacTel's intention to bid, which may explain why they failed to conduct proper investigations and cautiously submitted weak bids.

This paper proposes a simple model of a second-price auction in which buyers can acquire costly information before bidding. Buyers' valuations are drawn i.i.d. from some common knowledge distribution, but are unknown to buyers ex-ante. The main innovation of our model is that it gives buyers flexibility in *what* information they can seek. Specifically, buyers can acquire information about both their own valuations for

[1] Due diligence for the acquisition of a company can take several months and involves high legal and advising fees.

[2] For instance, the Chicago license was sold at $31 per capita. See https://www.fcc.gov/auctions-summary for the auction outcome and Klemperer (2002) for a more detailed account.

the good as well as those of their competitors.[3]

We model information acquisition as a two-step process. Buyers first have the opportunity to assess the potential competition by acquiring a signal about other bidders' valuations. Then, after having observed that signal's realization, they decide what signal to acquire about their own values. Buyers have some flexibility in choosing a signal and, in particular, can choose how the 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 never converge to becoming fully informed of their valuations in equilibrium, even as information costs become arbitrarily small relative to the value of the good. 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 when striving for the same good, as depicted in Figure 1.

The rest of the paper examines 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 when the cost of information is small. Indeed, losing buyers often fail to learn their valuations precisely, and since they bid 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]

[3] Learning about the competitors can be interpreted as undertaking market research. For instance, bidders in broadband subsidies auctions often start by investigating how much the subsidy is worth to a company with a different technology than their own—e.g., how much it is worth to a company providing broadband via satellite, whereas they use cable. The sole purpose of such information is to predict others' bids and assess their chances of winning the auction.

[4] This effect persists even as information costs become arbitrarily small. This highlights a discontinuity between the standard model, where buyers know their valuations ex-ante (the cost of information is

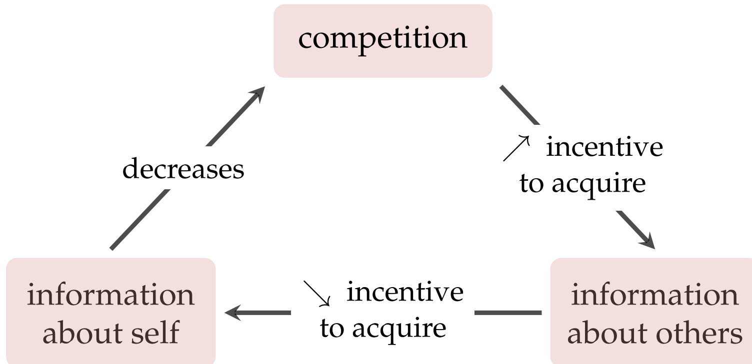


Figure 1: *The interplay between competition and information.*

Our first results highlight a new adverse effect of competition on revenue, and we investigate its implications for auction design. We show that attracting an extra bidder is often less valuable than setting an optimal reserve price,[5] suggesting that the seminal result of Bulow and Klemperer (1996) relies on buyers knowing their valuations fully. There are several forces at play. First, the presence of an additional bidder does not raise revenue as much as in the standard model, as it negatively impacts others' learning incentives. 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. Furthermore, since the optimal reserve price seeks to address this perverse effect of competition on revenue, it varies with the number of buyers N and converges to the highest possible valuation as N grows large. This contrasts with the standard model in which the optimal reserve price is independent of N.

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 maintains uncertainty over the extent of competition. Buyers can no longer predict whose bids they will be facing in the auction, which reduces their 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 unambigu-

then effectively zero), and ours.

[5]Specifically, we show that this is the case whenever the number of bidders N is not too small. A reserve price is a minimum acceptable bid set by the seller, which is allowed to depend on the primitives of the model (i.e., the number of buyers and the distribution of buyers' valuations).

ously 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.[6]

Finally, we investigate whether auctions still perform their role of price discovery once information is endogenous and shaped by competition. By running an auction, the seller not only raises revenue but also elicits information about buyers' valuations. That information might be valuable in itself—e.g., it might help the seller assess the demand for his good, which is useful for making production decisions or setting a reserve price. To that end, we extend our baseline model and introduce a common component that determines the distribution of buyers' valuations. When information is exogenous, buyers' bids reflect the entire distribution of valuations, thus allowing the seller to learn the common component as the number of buyers N gets large. This is no longer the case in our setting: because buyers only learn their valuations when they have a chance of winning, their bids are much less informative and can fail to reveal the common component. In particular, when the number of buyers is large, competition almost always occurs at the upper tail of the distribution of valuations. In equilibrium, buyers then only learn their valuations when it falls into that range, and their bids are only informative of that tail of the distribution. Not only are bids less informative, but the auction price (i.e., the second-highest bid) is also less likely to reflect the good's value (i.e., the highest valuation). Indeed, the price oftentimes lies strictly below the second-highest valuation and, as a result, converges to the good's value much more slowly than in the standard model.

Overall, our results suggest that learning incentives undermine the two main benefits of competition in auctions: increasing revenue and price discovery.

[6]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 and foremost, we build on a previous paper (Gleyze and Pernoud (2022)), which considers a general mechanism design setting in which participants can acquire costly information on their preferences as well as others'. We show that most selling mechanisms incentive participants to learn about others' preferences. Hence, such an incentive cannot be designed away by the seller and should be expected in most settings. In this paper, we investigate how that affects the value of competition in auctions. 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.

Second, our paper speaks to the literature highlighting the value of competition in selling mechanisms, which often compares auctions to "negotiations." Bulow and Klemperer (1996) argue that auctions have the benefit of attracting more buyers at the cost of less negotiating power for the seller,[7] but show that attracting just one more buyer has more value than being able to commit to the 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.[8] These results, however, take buyers' information as fixed. Instead, our paper asks how competition affects the information buyers acquire and reaches opposite conclusions.

A growing empirical literature investigates why some sellers favor negotiations over auctions. Most of it focuses on the market for corporate control, in which both types of selling mechanisms are commonly observed (Boone and Mulherin (2007); Aktas et al. (2010); Gentry and Stroup (2019)).[9] Takeover auctions are particularly relevant applications for our analysis as they involve high due diligence costs. The same is true for procurement auctions, and Bajari et al. (2009) find that auctions tend to perform poorly when procurement projects are complex, which is consistent with our results.

Our paper also contributes to the vast literature on entry and learning costs in auctions. Levin and Smith (1994) characterize the symmetric (mixed) equilibrium under a second-price auction when buyers pay a fixed cost to learn their values before bidding.

[7] E.g., the seller might no longer be able to withdraw the object for sale if bids are too low.

[8] 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.

[9] Boone and Mulherin (2007) find that about half of the targets of corporate takeovers in the U.S. in the 1990s were auctioned between multiple buyers, while the remainder negotiated with a single buyer.

They show that equilibrium entry decisions are efficient and revenue-maximizing in the IPV setting.[10] In a similar setting, Compte and Jehiel (2007) show that dynamic auction formats tend to dominate static ones, as dynamic formats reveal more information on the toughness of competition. Another strand of the literature on learning costs 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 only allow buyers to learn about their own valuations. Instead, our paper seeks to understand how competition affects buyers' learning incentives. To that end, we allow buyers to also learn about their competitors' valuations, leading to new insights and different predictions. One contribution of our paper is then to develop a tractable model of *multidimensional* learning in auctions.

A relatively small literature studies buyers' incentives to learn about their competitors' types in first-price auctions (Tian and Xiao (2007)) and auctions with interdependent values (Bobkova (2019); Kim and Koh (2020)).[11] Information about opponents' types is valuable as it allows buyers to either alleviate the winner's curse (in interdependent-value auctions) or shade their bids more aggressively (in first-price auctions). Such incentives are absent in our setting as buyers compete in a second-price auction and their valuations are independent and private. Buyers then only learn about the competitors so as to assess how much they should learn about themselves.

Finally, several papers study the performance of auctions when buyers are ex-ante asymmetric (Maskin and Riley (2000); Kim and Che (2004); Cantillon (2008)). Jehiel and Lamy (2015) investigate optimal auction design with asymmetric buyers and endogenous entry, and show that the designer should discriminate against incumbents but not against entrants. Marquez (2021) shows that the presence of an asymmetrically strong buyer hurts revenue if it deters entry of remaining "regular" buyers. Compte and Jehiel (2002) show that more competition in the auction can reduce total welfare if buyers are asymmetrically informed and have interdependent valuations. In our paper, buyers are ex-ante identical, but we show that asymmetries in private information arise endogenously, even in symmetric equilibria.

[10]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.

[11]In these papers, learning is one-dimensional: buyers usually know their own types and acquire costly information about their opponents'. In Bobkova (2019), buyers choose whether to learn about a private versus a common component in their valuations, but they cannot do both.

2 The Model

A seller puts a unique, indivisible good for sale through a second-price auction.[12] 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.[13] 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.

[12] 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.

[13] 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 $v = (v_i)_i$ at some cost before competing in the auction. They also learn their own (and only their own) noise terms 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 generality, buyers first acquire information about others' valuations and, conditional on the realization of this signal, acquire information on their own. Indeed, since we are looking at a (strategy-proof) second-price auction, information about others is only valuable insofar as it helps buyers decide how much they should learn about themselves. 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 $\tilde{v}_{-i}$.[14]

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 *convex* 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 convex 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 $\Pi_i^{other} \in \Delta\mathcal{P}$ (buyers are allowed to randomize over information partitions). Then,

[14]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.

conditional on his information set about $\max_j v_j$, he chooses an information partition about his own valuation $\Pi_i^{self} : 2^V \longrightarrow \Delta\mathcal{P}$.

Finally, each buyer chooses a measurable bidding strategy $\sigma_i : 2^V \times 2^V \times U \longrightarrow \mathbb{R}_+$, which outputs a bid given the buyer's overall information $\pi_i = (\pi_i^{other}, \pi_i^{self}, u_i)$.

Buyer i's ex-ante expected utility under strategy profile $(\Pi_i^{other}, \Pi_i^{self}, \sigma_i)_i$ writes[15]

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

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 $(\Pi_i^{other}, \Pi_i^{self}, \sigma_i)$ 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. Intuitively, we require that each buyer's equilibrium strategy remains optimal when his opponents might tremble with vanishing probability and make a bid drawn from a full-support distribution. Formally, a Nash equilibrium $(\Pi_i^{other}, \Pi_i^{self}, \sigma_i)_i$ is robust to trembles if there exist a distribution F with support $[\min_{\nu_i \in V\times U} \nu_i, \max_{\nu_i \in V\times U} \nu_i]$ and a sequence of positive numbers $\{\varepsilon^k\}_{k=1}^{\infty}$ converging to zero such that, for all buyers i and equilibrium information sets π_i, $\sigma_i(\pi_i)$ is a best response to $(\Pi_j^{other}, \Pi_j^{self}, \hat{\sigma}_j^k)_{j\neq i}$, where $\hat{\sigma}_j^k$ is a perturbed bidding strategy for j that equals σ_j with probability $1 - \varepsilon^k$ and F otherwise.

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 $\Pi = \{\pi_l\}_{l=1,\ldots L}$ induces a distribution over posterior beliefs, which puts weight

[15] We extend the domain of c to mixed strategies in the usual way, with $c(\Pi, \cdot)$ denoting the expected cost of information partition Π.

on as many posteriors as there are elements in the partition $\{\mu_l\}_{l=1,\dots L}$ with

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

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.

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 $H(p) = -\sum_v p(v) \log[p(v)]$.*[16] *The cost of an information partition equals the expected reduction in the entropy of the buyer's belief that it induces:*

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

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.

[16] 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 as $\hat{p} \longrightarrow p$, for some full-support $\hat{p}$, of $-\sum_v \hat{p}(v) \log[\hat{p}(v)]$.

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 elements: 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 $\Pi^{self}_{\leq v^*} \equiv \{\{v\}_{v \leq v^*}, \{v : v > v^*\}\}$. Hence even a coarse signal about others can significantly reduce how finely a buyer should learn about his own valuation.

Lemma 1. *There exists $\overline{\Sigma}$ and $\overline{m}$ such that if $\sum_v [p(v)]^2 \leq \overline{\Sigma}$ and $m \geq \overline{m}$, then it is cost efficient for buyers to acquire* some *information about others first. Formally,*

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

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 always lies weakly below one, is equal to one only if the prior is deterministic (i.e., $p(v) = 1$ 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 valuations 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.

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 not too large, buyers must choose the fully revealing partition $\Pi^{self} = \{\{v\}_{v \in V}\}$.

3.2 The General Case

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 first 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 2. *Let $(\Pi_\lambda^{other}, \Pi_\lambda^{self}, \sigma_\lambda)$ by any equilibrium given cost parameter $\lambda > 0$. There exists $\varepsilon > 0$ such that*

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

To 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. Indeed, buyer j might have learned something about i's valuation, in which case j's bid might carry information that is relevant to i. 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., $\sigma_j(\pi_j) = v_i + \underline{u}$). 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.

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 (alternatively, auctions with small cost parameter λ) and symmetric equilibria enables us to characterize equilibrium information structures.[17]

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, in any symmetric equilibrium, if an information structure has probability $\Pr\left(\Pi_\lambda^{other}, \Pi_\lambda^{self}\right) \geq \varepsilon(\lambda)$, then it must satisfy:*

$$\Pi_\lambda^{other} \neq \{V\} \text{ and } \Pi_\lambda^{self}(\pi^{other}) = \left\{\{v_i \mid v_i < \min_{v \in \pi^{other}} v\}, \{v_i\}_{v_i \in \pi^{other}}, \{v_i \mid v_i > \max_{v \in \pi^{other}} v\}\right\}$$

for all $\pi^{other} \in \Pi^{other}$.

In words, buyers acquire *some* information about others in equilibrium, and then only learn their own valuations if they fall in a similar range as that of their toughest competitor. In particular, they fail to distinguish all valuations at which they lose the auction for sure and, similarly, all valuations at which they win for sure. Figure 2 depicts an equilibrium information structure. Overall, the competitive pressure that buyers impose on each other shapes the information that buyers acquire in significant ways. The rest of the paper examines how that, in turn, affects the value of competition.

[17]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.

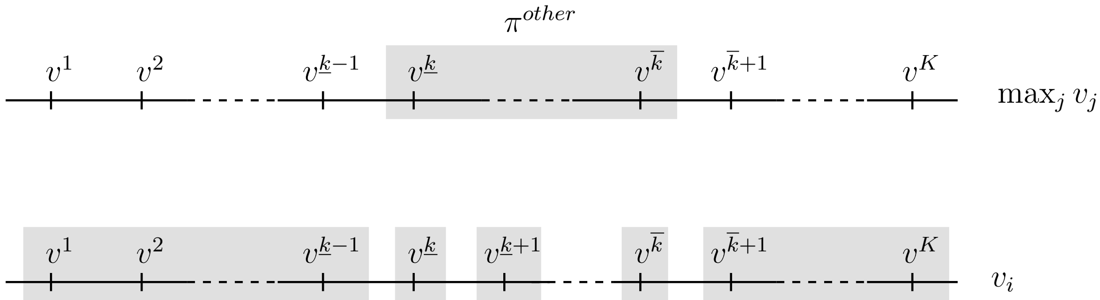


Figure 2: *Let* $V = \{v^1, v^2, \ldots, v^K\}$ *and order the valuations in increasing order, i.e.,* $v^k < v^{k+1}$. *If buyer* i *learns that* $\max_j v_j \in \pi^{other} \equiv \{v : v^{\underline{k}} \leq v \leq v^{\bar{k}}\}$ *(top line), he chooses to fully learn his valuation if it belongs to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

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* $\overline{\lambda} > 0$ *such that, for all* $\lambda \leq \overline{\lambda}$*, the revenue generated in any symmetric equilibrium* $\left(\Pi_\lambda^{other}, \Pi_\lambda^{self}, \sigma_\lambda\right)$ *is bounded away by* L *from the expected second-highest valuation:*

$$\mathbb{E}\left[\text{ equilibrium revenue } \mid \left(\Pi_\lambda^{other}, \Pi_\lambda^{self}, \sigma_\lambda\right)\right] < \mathbb{E}\left[\nu_{(2)}\right] - L.$$

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.

The intuition is simple. 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. 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 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 $V = \{\frac{1}{K}, \frac{2}{K}, \ldots, \frac{K-1}{K}, \frac{K}{K}\}$ 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 information structures as described in Figure 2. Under such information structures, buyers must acquire some information about others $\Pi^{other} \neq \{V\}$, and only learn to distinguish their own valuations when they fall in the same range as that of their toughest competitor. We consider the symmetric tremble-robust equilibrium constructed in Section B.2.3 of the Appendix, in which buyers choose the cheapest such information structure. We find this information structure numerically, and depict the equilibrium information partition about others Π^{other} for several values of N in Figure 9 of Appendix A.1. We then simulate equilibrium bids and compute expected revenue for vanishing λ.[18] We also compute expected revenue when buyers are fully informed of their valuations, which equals the expected second-highest valuation.

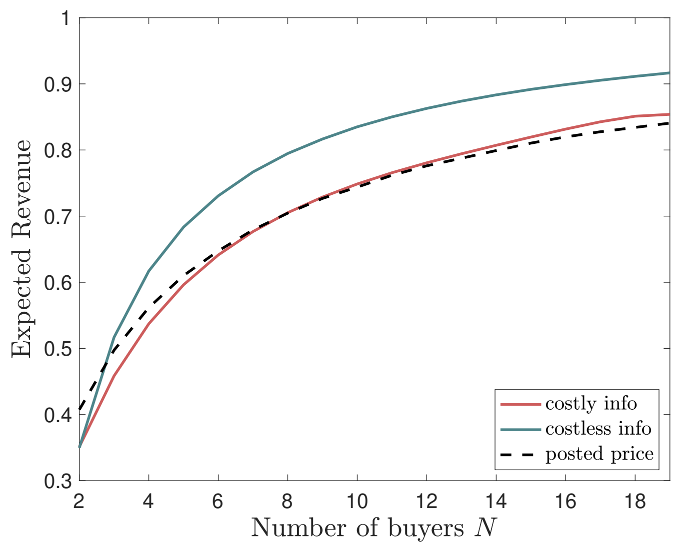
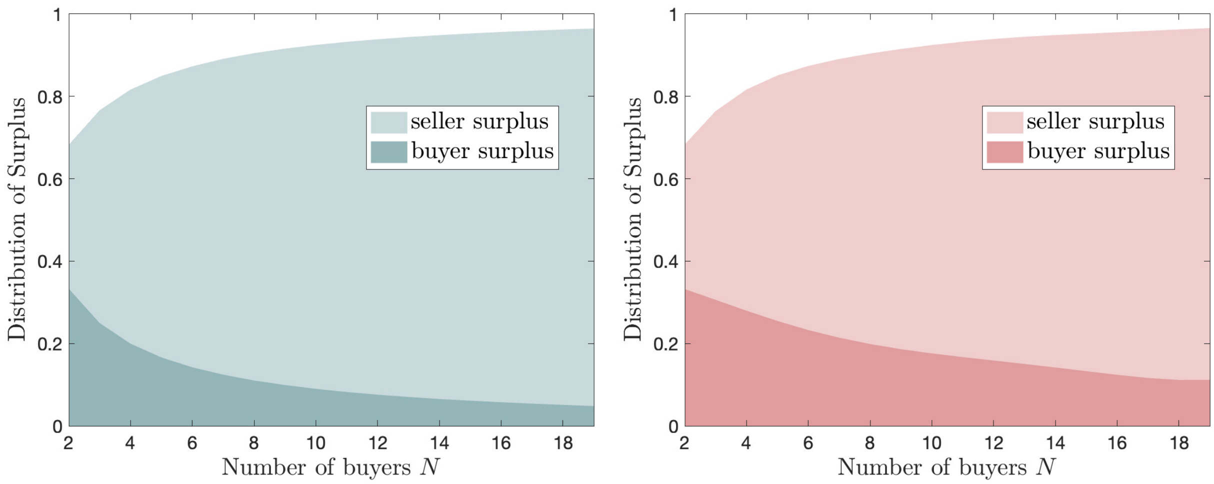


Figure 3: *The top (blue) line plots expected revenue in the standard model (i.e., the expected second-highest value). The bottom (red) line plots expected revenue in our model for small λ, fitted with a 6^{th}-degree polynomial to smooth small irregularities arising from the discreteness of V. For comparison, 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 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.[19] 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).

[18] 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).

[19] 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.



(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* $K = 30$.

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 $\sigma_i(\pi_i) > 0$.

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 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, that is:[20]

$$\Pr\left(\max_j \nu_j < \nu^*\right) \nu^* = \kappa.$$

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 a symmetric 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 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 symmetric equilibrium $\left(\Pi_\lambda^{other}, \Pi_\lambda^{self}, \sigma_\lambda\right)$, the probability that at least two buyers enter the auction is bounded above by*

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

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

Proposition 3 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. 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.[21]

[20] If only one buyer enters the auction, that buyer wins at zero price.

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

A direct corollary of Proposition 3 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 3), 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 randomizing access to the auction. Doing so maintains uncertainty over the extent of competition and incentivizes buyers to learn their valuations. Overall, this suggests that competition is most effective if it is carefully *designed* by the seller.

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.

Two facts bring some nuance to this common wisdom. First, many high-value sales operate via negotiations with just a few buyers instead of open auctions. Indeed, about half of firm acquisitions occur through negotiations instead of auctions (Boone and Mulherin (2007)). Second, and as mentioned in the introduction, auctions with many bidders sometimes fail. This suggests that attracting additional participants may not always be as valuable as previously thought.

We show that, once we account for how competition shapes learning incentives, BK's theorem may no longer hold.

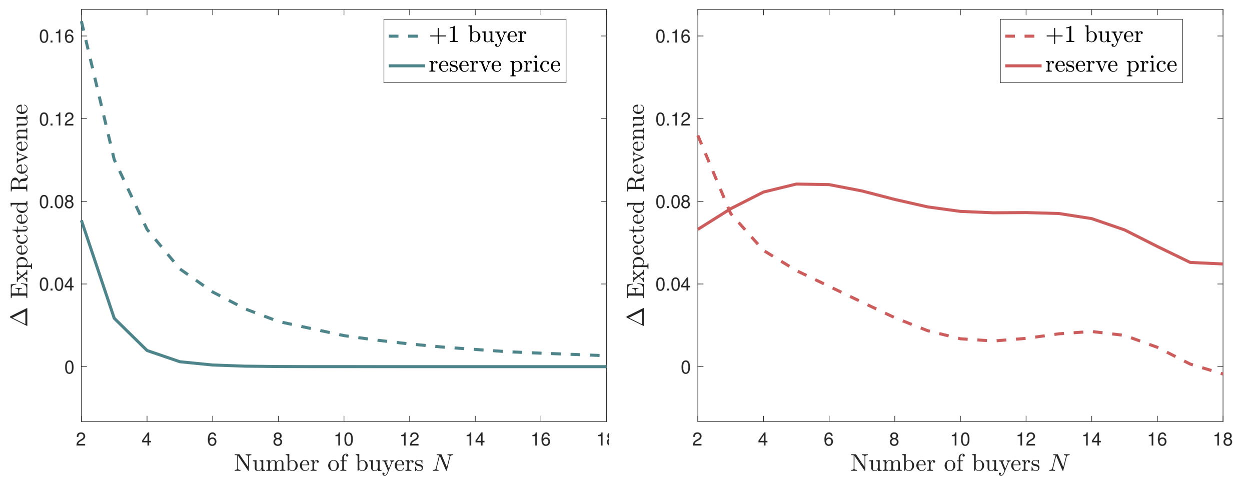
Theorem 3. *There exists $\overline{N}$ such that, for all $N \geq \overline{N}$ and for λ small enough, revenue with $N+1$ bidders in the second-price auction without reserve is **lower** than revenue with N bidders in the second-price auction with optimal reserve.*

That are several forces underlying Theorem 3. First, competition is less valuable than in the standard model, as additional buyers not only fail to learn their valuations fully but also exert a negative externality on others' learning incentives. For some values of N, it can be that adding a buyer to the auction changes the equilibrium information structure, leading buyers to learn more about their competitors—a finer Π^{other}—and less about themselves—a coarser Π^{self} in expectation. This has an adverse effect on expected revenue. 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. In the proof, we show that these forces favor a reserve price whenever the number of buyers N is sufficiently large. Our uniform example suggests that N does not have to be very large for Theorem 3 to apply.

Uniform Example Continued. We revisit the uniform example from Section 3.3 to illustrate Theorem 3. 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)).[22] In our framework, the optimal

[22]This result technically requires buyers' set of possible valuations to be an interval and the distri-



(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.*

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 10 in Appendix A.1 for an illustration.

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 $N = \{1, \dots N\}$, 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.

---

bution to satisfy an appropriate regulatory condition. It holds approximately when our finite set of valuations is sufficiently fine.

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 $\tilde{M}$. For completeness, suppose that if only one buyer is allowed access to the auction $|M| = 1$, then the buyer is offered the good at a (potentially random) posted price.

Proposition 4. *Take any $\varepsilon > 0$. There exists an access rule $\tilde{M}$ such that, for λ small enough,*

$$\mathbb{E}\left[\textit{equilibrium revenue} \mid \left(\Pi_\lambda^{other}, \Pi_\lambda^{self}, \sigma_\lambda\right)\right] \geq \mathbb{E}\left[\nu_{(2)}\right] - \varepsilon$$

in any symmetric equilibrium $\left(\Pi_\lambda^{other}, \Pi_\lambda^{self}, \sigma_\lambda\right)$.

Proposition 4 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 4, 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.

For instance, suppose the seller randomizes access in the following way. With probability $q < 1$, all buyers get access to the auction $M = \{1, \ldots, N\}$. All bids are then taken into consideration: the good goes to the highest bidder who pays the second-highest bid. With probability $1 - q$, 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 $(1 - q)/N$ 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.

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.

Theorem 4. *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. Theorem 4 then says that in our framework, there is more value in incentivizing buyers to acquire information than in reducing their information rent. We revisit one last time our uniform example in Figure 6 to illustrate this.

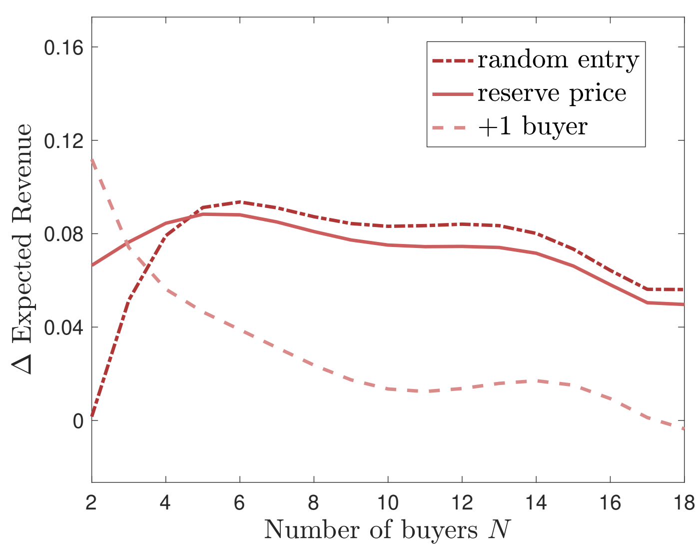


Figure 6: *The dash-dotted line plots the value of randomizing access to the auction—i.e., the difference between expected revenue when access is randomized optimally and expected revenue when all buyers are allowed into the auction—for λ arbitrarily small. The other two lines are the same as in Figure 4, and represent the value of a reserve price and of an additional buyer. 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.*

6 Failures of Price Discovery

In this section, we show that learning incentives undermine auctions' role of price discovery. Indeed, losing buyers often fail to learn their valuations precisely, which has two implications. First, the equilibrium distribution of bids carries much less information than in the standard model. Second, the equilibrium price converges to the good's "true" value (i.e., the highest valuation among buyers) at a slower rate.

6.1 Learning from Bids

In addition to raising revenue, auctions serve a role of price discovery. By running an auction, the seller collects information from all the buyers, which can then be used not only to find a correct price but also to learn about the good's underlying value more broadly. For instance, buyers' bids provide information on the demand for the good (i.e., on the distribution from which buyers' valuations are drawn), which is of interest to the seller as it is a key determinant of the optimal reserve price. In practice, sellers often use past auction data to set reserve prices—this is for instance the case in Canadian timber auctions (Athey et al. (2003)) and online advertising (Ostrovsky and Schwarz (2011)).

To assess this role of price discovery, we extend our model and introduce a common component $\omega \in \Omega$ that affects the distribution of buyers' valuations. The common component is drawn from a finite set $|\Omega| < \infty$ according to some prior distribution $\mu_0 \in \Delta\Omega$. Given a realization of ω, buyers' valuations are independently drawn from the same full-support distribution $p_\omega \in \Delta V$. The seller gathers information about the common component ω by running an auction. Indeed, the seller gets to observe buyers' bids, which are informative of buyers' valuations, and hence of ω. We investigate whether the auction reveals ω as the market grows large.

Let $\left(\Pi_N^{other}, \Pi_N^{self}, \sigma_N\right)$ be a symmetric tremble-robust equilibrium of the auction when N buyers participate in it.[23] An equilibrium induces a distribution over vectors of bids, and, given realized bids, the seller updates his belief about ω using Bayes' rule. Let $\tilde{\mu}_N$ be the seller's (random) posterior about the common component ω. If the seller's posterior always equals the prior $\Pr(\tilde{\mu}_N = \mu_0) = 1$, then the seller does not learn anything by running the auction. On the contrary, if $\Pr(\tilde{\mu}_N = \delta_\omega) = \mu_0(\omega)$

[23] The equilibrium still depends on information costs λ, but we drop the indexing to keep the notation uncluttered.

for each $\omega \in \Omega$, where δ_ω is the corner belief that puts probability one on the common component being ω, the distribution of bids fully reveals the common component.

Definition 1. *A sequence of auctions* $\left(\Pi_N^{other}, \Pi_N^{self}, \sigma_N\right)$ *reveals the common component if the seller's posterior* $\tilde{\mu}_N$ *converges in probability to* $\delta_{\tilde{\omega}}$ *as* $N \longrightarrow \infty$.

Proposition 5. *Suppose buyers can only learn about themselves, so* $\Pi_N^{other} = \{V\}$. *For each* N, *if the cost parameter* λ *is sufficiently small, buyers become fully informed about their valuations and bid truthfully. The sequence of auctions then reveals the common component as* $N \longrightarrow \infty$ *whenever* $p_\omega \neq p_{\omega'}$ *for all* $\omega \neq \omega'$.

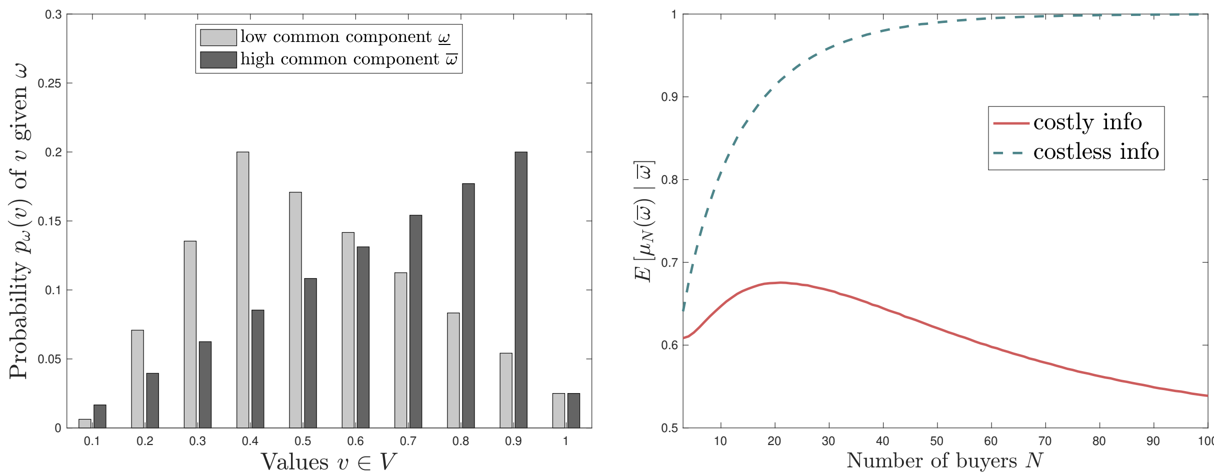
If buyers can also learn about their competitors, the equilibrium distribution of bids is, however, much less informative of the common component ω. Indeed, we know from the above analysis that, in equilibrium, buyers first assess the toughness of the competition, and then only learn their valuations if they fall in a similar range as their toughest competitor's. Bids are then only informative in that (potentially small) range of valuations in which competition is intense and buyers learn their valuations. When N is large, competition is very likely to be at the very top of the distribution, and equilibrium bids are only informative of that tail of the distribution. This might not be enough to infer the common component ω—e.g., if $p_{\omega'}$ and $p_{\omega''}$ coincide at the very tail of the distribution for some $\omega' \neq \omega''$.

Proposition 6. *Now consider our main model specification in which buyers can also learn about their competitors. There exist environments* $\{p_\omega\}_\omega$ *with* $p_\omega \neq p_{\omega'}$ *for all* $\omega \neq \omega'$ *such that the sequence of auctions* ***does not*** *reveal the common component as* $N \longrightarrow \infty$, *even when the cost parameter* λ *is arbitrarily small.*

The following example illustrates Proposition 6. The set of possible valuations is $V = \left\{\frac{1}{K}, \ldots, \frac{K-1}{K}, \frac{K}{K}\right\}$ with $K = 10$, and there are two possible realizations of the common component $\Omega = \{\underline{\omega}, \overline{\omega}\}$, which are equally likely ex-ante. The distributions of valuations $(p_{\underline{\omega}}, p_{\overline{\omega}})$ are depicted in Figure 7 (left panel). Intuitively, higher values are more likely when the common component is high—the mode of the distribution is higher when $\omega = \overline{\omega}$—but not uniformly so: both distributions are "bell-shaped" and look similar for very high realizations of v. In particular, they assign the same probability to the highest possible valuation $v = 1$.

As in the uniform example, we consider the equilibrium constructed in Section B.2.3 of the Appendix and set λ to be arbitrarily small.[24] We compare how fast the

[24]We set the measure of uncertainty H to be the entropy.



(a) Distributions of values $p_{\underline{\omega}}$ and $p_{\overline{\omega}}$

(b) Seller's learning

Figure 7: *The right panel plots the expected probability that the seller's posterior belief puts on $\omega = \overline{\omega}$ when the realized common component is indeed high. The seller learns the common component if and only if that probability goes to one as the number of buyers N grows. In this example, he does not.*

seller learns about ω in the standard model and in our model with endogenous information using numerical simulations. To do so, we assume that the realized common component is $\omega = \overline{\omega}$, draw buyers' valuations from $p_{\overline{\omega}}$, and simulate equilibrium bids. Each simulation leads the seller to hold some posterior belief μ_N about ω. In Figure 7 (right panel) we plot the expected weight he puts on the high (and hence correct) realization of the common component $\mathbb{E}[\mu_N(\overline{\omega}) \mid \overline{\omega}]$.

When buyers know their valuations and bid truthfully, the seller quickly learns that $\omega = \overline{\omega}$: his expected posterior quickly converges to one (dashed blue line). This is however not the case in our model: the seller fails to learn the common component as N grows large (the solid red line on the right panel never goes to one). Indeed, for N large enough, buyers find it optimal to learn whether their toughest competitor has the highest possible valuation, as this event is very likely. If this is the case, they then only learn whether their own valuation is the highest one possible as well or whether it is lower. Their bids are then only informative of whether they have the highest valuation possible, which is not informative of ω since $p_{\overline{\omega}}(1) = p_{\underline{\omega}}(1)$.[25] Furthermore,

[25]Of course, this example is extreme as it assumes the tail of p_ω to be independent of ω. More generally,

the seller's learning is non-monotonic in the number of buyers N and maximal for intermediate values of N. Indeed, when there are only very few buyers, there are only very few draws of valuations to learn from. When the number of buyers is large, intense competition crowds out learning incentives, and bids are only informative of the very tail of the distribution, which is here independent of the common component. Overall, in this example, the auction is most informative with an intermediate number of buyers N.

6.2 Slower Price Convergence

The previous section examined how much information about the good is revealed by buyers' bids. A related but distinct question asks whether the *price* appropriately reflects the good's "true" value.[26] This question has received much attention in the literature as it not only provides a theoretical foundation for price formation but also motivates the use of auctions to set prices in practice. In many real-world markets, and particularly markets for commodities such as electricity or timber, auction prices are used as benchmarks in many long-term transactions. For instance, in a typical wholesale electricity market, only a small percentage (ranging from 5% to 20%) of the electricity volume is traded on the auction-based spot market. The auction outcome however matters for the entire market, as it is used to price all remaining long-term trades. It is then all the more important that the auction price be "correct".

In our model, the good's "value" coincides with the highest valuation among the buyers $\nu_{(1)}$. The price then appropriately reflects the good's value if it is sufficiently close to $\nu_{(1)}$. In the standard auction model, the competitive pressure between buyers reduces their information rent and fosters information aggregation. Several papers have indeed shown that, as the number of buyers gets large, the auction price converges to the true value of the good for sale (Wilson (1977); Milgrom (1979), etc.). This is still true in our framework, but convergence is much slower. Indeed, when losing buyers fail to learn their valuations, they bid less aggressively and leave a greater gap between the second-highest bid (the good's *price*) and the highest valuation (the good's

learning incentives impede price discovery whenever the tail of the distribution of buyers' valuations is not informative enough of the underlying parameter of interest. Even when the auction does reveal ω as N grows large, it does so at a much slower rate. We illustrate this in Section A.2 of the Appendix.

[26]Formally, in the former section, the seller learns about some underlying parameter of interest by observing the bids of *all* buyers. This section focuses on whether the second-highest bid *by itself* correctly reflects the good's value.

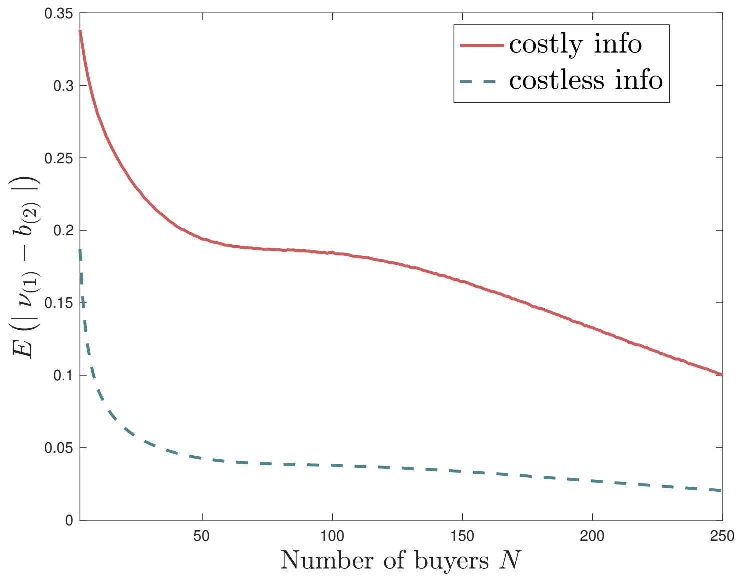


Figure 8: *The lines plot the expected difference between the highest valuation and the equilibrium price in the standard model (dashed blue line) and in our model for small λ (solid red line). The set up is the same as in Figure 11.*

value). In Figure 8 we revisit the example from the previous section and plot the expected gap between the good's value and its price as a function of N. This gap goes to zero at a much slower rate once we account for learning incentives.

Slow convergence can be a concern in practice: it suggests that a larger volume of trades needs to be executed via auctions for the resulting price to be reliable enough. In settings where learning incentives are of first-order importance, the standard auction model might underestimate how large an auction is needed to find an appropriate price.

7 Conclusion

This paper develops a model of information acquisition in auctions, in which buyers can first learn about the strength of competition before learning about their own valuations. We first characterize how the competitive pressure between 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 hurts revenue and price discovery. We propose market design solutions to mitigate these effects. Overall, we show that the seller benefits from carefully designing the competition that buyers face—either by imposing it himself 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 unknown 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

Aktas, N., De Bodt, E., and Roll, R. (2010). Negotiations under the threat of an auction. *Journal of Financial Economics*, 98(2):241–255.

Athey, S., Cramton, P., and Ingraham, A. (2003). Upset pricing in auction markets: An overview. *White Paper, Market Design Inc., on behalf of British Columbia Ministry of Forests*.

Bajari, P., McMillan, R., and Tadelis, S. (2009). Auctions versus negotiations in procurement: an empirical analysis. *The Journal of Law, Economics, & Organization*, 25(2):372–399.

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.

Boone, A. L. and Mulherin, J. H. (2007). How are firms sold? *The Journal of Finance*, 62(2):847–875.

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.

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.

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.

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

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.

Milgrom, P. R. (1979). A convergence theorem for competitive bidding with differential information. *Econometrica: Journal of the Econometric Society*, pages 679–688.

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

Ostrovsky, M. and Schwarz, M. (2011). Reserve prices in internet advertising auctions: A field experiment. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 59–60.

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

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

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

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.

Wilson, R. (1977). A bidding model of perfect competition. *The Review of Economic Studies*, 44(3):511–518.

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 consider the equilibrium constructed in Section B.2.3, in which all buyers choose the cost-minimizing information structure $\left(\Pi^{other}, \Pi^{self}\right)$ 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 9. 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 $.72$.

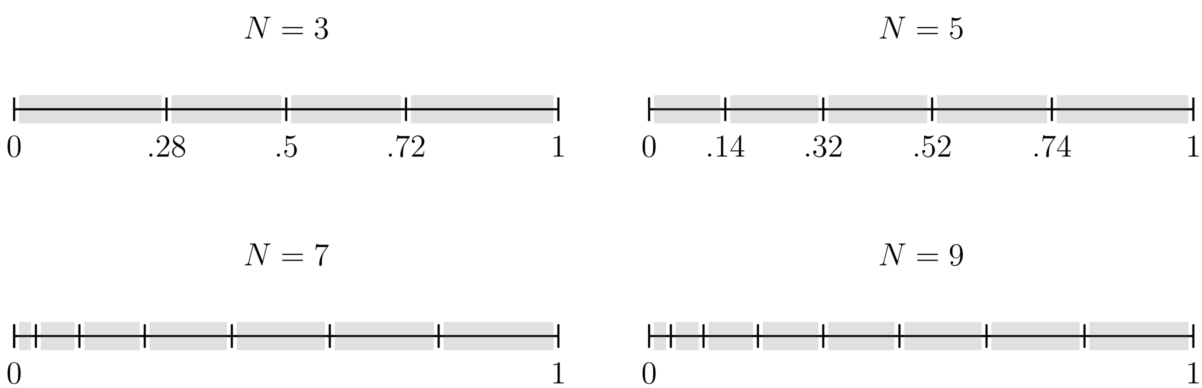


Figure 9: *Parameter* $K = 50$.

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 10. 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).

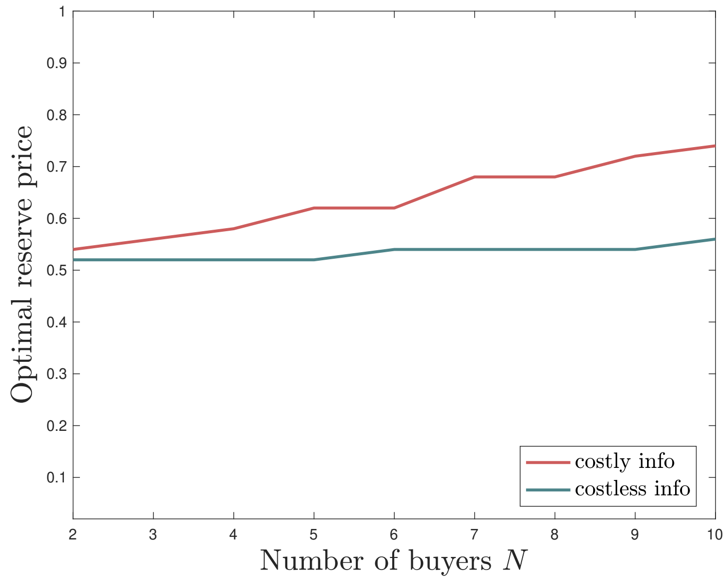
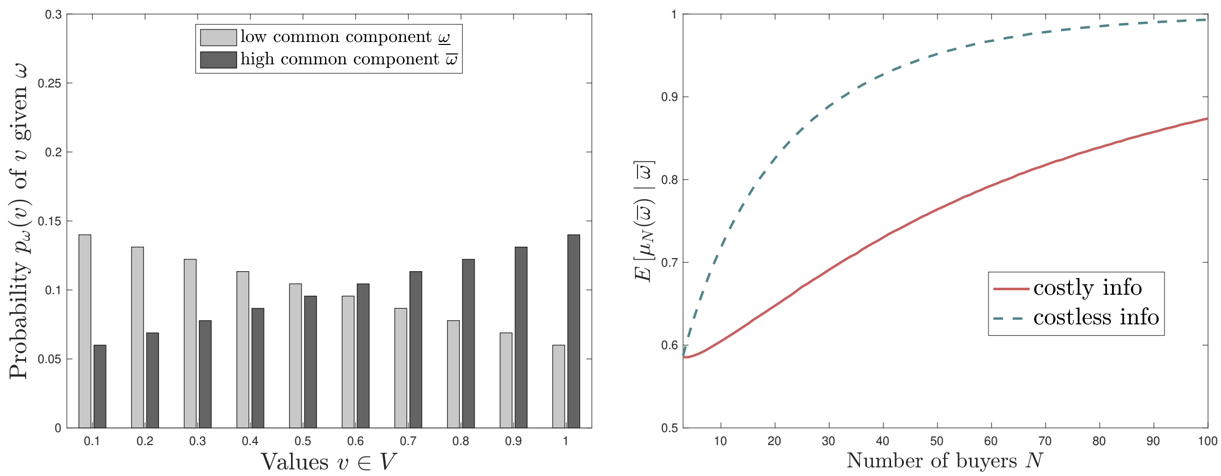


Figure 10: *The optimal reserve price under exogenous information (blue line) is close to being independent of N. The fact that it varies slightly with N comes from the discreteness of V, and disappears for K large enough. On the contrary, the optimal reserve price when information is endogenous and shaped by competition is increasing in N. Parameter $K = 50$.*

A.2 Failure of Price Discovery: Another Example

We now give an example that illustrates how endogenous information acquisition slows down the rate at which the seller learns the common component. Hence, even when the auction reveals ω, buyers' learning incentives still significantly affect price discovery.

Consider the distributions $(p_{\underline{\omega}}, p_{\overline{\omega}})$ depicted in Figure 11 (left panel). When the common component is low $\omega = \underline{\omega}$, lower values are uniformly more likely (i.e., $p_{\underline{\omega}}(v) > p_{\underline{\omega}}(v')$ whenever $v < v'$). The reverse is true when the common component is high $\omega = \overline{\omega}$. We solve for the equilibrium, simulate equilibrium behavior assuming $\omega = \overline{\omega}$, and plot the seller's posterior in Figure 11 (right panel). When buyers know their valuations and bid truthfully, the seller quickly learns that $\omega = \overline{\omega}$: his expected posterior quickly converges to one (top blue line). His learning is much slower when information is endogenous (bottom red line).



(a) Distributions of values $p_{\underline{\omega}}$ and $p_{\overline{\omega}}$

(b) Seller's learning

Figure 11: *The right panel plots the expected probability that the seller's posterior belief puts on $\omega = \overline{\omega}$ when the realized common component is indeed high. The seller learns the common component if and only if that probability goes to one as the number of buyers N grows.*

A.3 Discussion and Robustness of the Model

The key assumptions in our analysis are the ones 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.

Proposition 7 (order of signals)**.** *Consider our main model specification, but suppose that buyers can now choose in which* order *to acquire the two signals.*[27] *All results remain unchanged.*

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. Proposition 7 says that such *ordering* is without loss within the class of two-step learning processes. Indeed, the only other alternative would be for buyers to first choose how much to learn about their own values before learning about others'. For sufficiently small information costs λ, they would then find it optimal to fully learn their values

[27] 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 the reverse.

and learn nothing about their competitors.[28] This, however, cannot be part of an equilibrium (Proposition 2).

A more 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 buyers learn about the competition in aggregate—e.g., by undertaking market research or by hiring a consulting company to inform them of the overall competition. 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.

Proposition 8 (learning about $\max_j v_j$)**.** *Consider our main model specification, but suppose that buyers can now only learn about each of their competitors independently, such that the cost of partition* Π^{other} *equals* $(N-1)c\left(\Pi^{other}, p\right)$*. There exists* $\underline{N}$ *such that, for all* $N \leq \underline{N}$*, Theorems 1 and 2 hold.*

If the number of competitors is not too large, then it remains cost-efficient for buyers to acquire some information about the competition and only learn their own values when worth it. As N increases, inquiring about the competition becomes increasingly costly, and above some threshold it is no longer cost-efficient. Hence, under such alternative modeling, our results persist for auctions of intermediate size but not for large auctions.

[28] Indeed, information about others is only valuable insofar as it helps buyers determine what information to acquire about their own values.

Appendix B Proofs

B.1 Preliminary Analysis

Proof of Lemma 1. Acquiring partition $\Pi^{other}_{v^*}$ leads agent i to hold one of two possible posterior beliefs about $\max_j v_j$. Either he learns $\max_j v_j \leq v^*$, in which case his posterior is

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

or he learns $\max_j v_j > v^*$, in which case his posterior is

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

We want to show that, if $\sum_v [p(v)]^2$ is high enough, there exists $v^* \in V$ such that

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

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

This rewrites as

(1)

$$H(p_{1:N-1}) - \Pr(\max_j v_j \leq v^*) H\left(\delta_{v_{1:N-1} \leq v^*}\right) - \Pr(\max_j v_j > v^*) H\left(\delta_{v_{1:N-1} > v^*}\right)$$
$$< \Pr\left(\max_j v_j \leq v^*\right) \Pr\left(v_i > v^*\right) \left[H\left(p\left(\cdot \mid v_i > v^*\right)\right) - \sum_{v > v^*} \frac{p(v)}{\Pr(v_i > v^*)} H(\delta_v)\right]$$

$$+ \Pr\left(\max_j v_j > v^*\right) \Pr\left(v_i \le v^*\right) \left[H\left(p\left(\cdot \mid v_i \le v^*\right)\right) - \sum_{v \le v^*} \frac{p(v)}{\Pr(v_i \le v^*)} H(\delta_v)\right].$$

We first bound below the reduction in information costs on self that is induced by learning whether the competitor has value below or above v^*. That is, we want to find a lower bound for the RHS of (1). Note that

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

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:

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

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

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

Using the strong concavity of H iteratively, we then get

$$\begin{aligned}
&H\left(p\left(\cdot \mid v_j \le v^*\right)\right) - \sum_{v \le v^*} \frac{p(v)}{\Pr(v_i \le v^*)} H(\delta_v) \\
&\ge \frac{m}{2} \frac{p(v^1)}{\Pr(v_i \le v^*)} \left[1 - \frac{p(v^1)}{\Pr(v_i \le v^*)}\right] ||\delta_{v^1} - \delta_{v^1 < v_i \le v^*}||^2 \\
&+ \frac{m}{2} \left[1 - \frac{p(v^1)}{\Pr(v_i \le v^*)}\right] \frac{p(v^2)}{\Pr(v^1 < v_i \le v^*)} \left[1 - \frac{p(v^2)}{\Pr(v^1 < v_i \le v^*)}\right] ||\delta_{v^2} - \delta_{v^2 < v_i \le v^*}||^2 \\
&+ \dots
\end{aligned}$$

where

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

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

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

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

$$= \frac{m}{2} - \frac{m}{2} \frac{1}{\Pr(v_i \leq v^*)} \sum_{v \leq v^*} [p(v)]^2.$$

Similarly,

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

We know find an upper bound for the cost of $\Pi_{v^*}^{other}$, that is for the LHS of (1). Using the same steps as above,

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

$$+ \frac{m}{2} \left[1 - \frac{1}{\Pr(\max_j v_j \leq v^*)} \sum_{v \leq v^*} [p_{1:N-1}(v)]^2 \right],$$

and similarly for $H\left(\delta_{v_{1:N-1} > v^*}\right)$. Then,

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

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

Note that $H(p_{1:N-1}) - \sum_v p_{1:N-1}(v) H(\delta_{1:N-1=v}) = c\left(\{\{v\}\}_v, p_{1:N-1}\right)$, which is bounded above by some $\bar{c} < \infty$ since the cost of information is bounded. Combining all the above inequalities, condition (1) holds if

$$\bar{c} < m\left[\frac{1}{2} + \Pr\left(\max_j v_j \leq v^*\right)\Pr\left(v_i > v^*\right) + \Pr\left(\max_j v_j > v^*\right)\Pr\left(v_i \leq v^*\right)\right]$$

$$- m\left[\sum_v [p_{1:N-1}(v)]^2 + \Pr\left(\max_j v_j > v^*\right)\sum_{v \leq v^*}(p(v))^2 + \Pr\left(\max_j v_j \leq v^*\right)\sum_{v > v^*}(p(v))^2\right].$$

Let $\Sigma \equiv \sum_v (p(v))^2$, which lies weakly below one. No value $v \in V$ can have prior probability greater than $\sqrt{\Sigma}$. Hence, for small Σ, 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 for sufficiently small Σ, and hence strictly positive. For m high enough, the inequality must then hold strictly. □

Lemma 2. *There exists $\overline{\Sigma}$ such that if $\sum_v [p(v)]^2 \leq \overline{\Sigma}$, 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 $v^* > v^1$.

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

(2)

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

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

$$< \Pr\left(v^2 \leq \max_j v_j \leq v^*\right)\left[\Pr(v_i > v^*)H(\delta_{>v^*}) + \sum_{v \leq v^*}\Pr(v_i = v)H(\delta_v)\right]$$
$$+ \Pr\left(\max_j v_j > v^*\right)\left[\Pr(v_i \leq v^*)H(\delta_{\leq v^*}) + \sum_{v > v^*}\Pr(v_i = v)H(\delta_v)\right]$$
$$- \Pr\left(\max_j v_j > v^1\right)\sum_{v}\Pr(v_i = v)H(\delta_v).$$

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 (2) holds if

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

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 if $\sum_{v > v^1}\left(\frac{p(v)}{1-p(v^1)}\right)^2$ is small 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 and 2

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 lambda goes to zero. We construct a symmetric equilibrium (Π^{self}, σ) that has such property.

Consider the following symmetric strategy profile:

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

Since buyers' valuations are independent and private, and buyers only learn about themselves for the purpose of Proposition 1, then it is a dominant strategy for them to bid their true expected valuation for the good given their information set. 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 $v^1 < v^2 < \cdots < v^K$. Abusing notation, let $p^k = p(v^k)$.

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 valuess 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 wouldn't 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

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

If $u_{j^*} \leq u_i + \mathbb{E}[v_i|v_i \in \{v^k, v^{k+1}\}] - v^k$, then i wins with 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 induce a gain in gross payoff of

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

Hence, as soon as $\Pr(u_{j^*} > u_i) > 0$, the 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.[29] 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.

$\square$

Proof of Proposition 2. By contradiction, suppose there exists an equilibrium in which buyers converge to becoming fully informed about their valuations. 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 $\Pi_{k^*}^{other} = \{\{v^1, \ldots, v^{k^*}\}, \{v^{k^*+1}, \ldots, v^K\}\}$. 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 $\Pi_{>k^*}^{self} \equiv \{\{v_i\}_{v_i > v^{k^*}}, \{v_i : v_i \leq v^{k^*}\}\}$. Intuitively, he does not learn to distinguish all the valuations below the threshold, as he most likely

[29]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.

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 bid strictly above $v^{k^*+1} + \underline{u}$ despite his realized value being lower. This can only be optimal for that competitor 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^*}. Such bundling can be optimal only if that buyer expects no non-vanishing cost from doing so, given others' equilibrium strategy. In particular, that buyer must believe that with a probability very close to one he will not face a bid in that interval, and so must have learned that i's valuation is even higher. 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. 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 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.

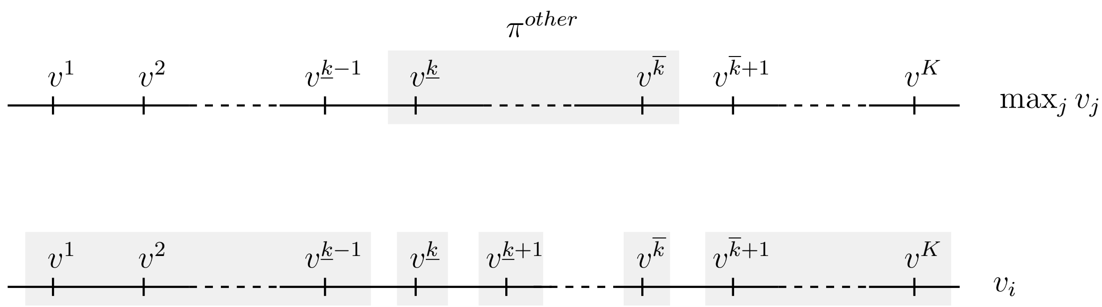
Under what conditions can the highest-value buyer j^*, with value $\nu_{j^*} > v^{k^*} + \overline{u}$, bid strictly below $v^{k^*} + \overline{u}$ in equilibrium? For j^* to find such low bid optimal, it has to be that buyer j^* fails to learn his value: that is, he sometimes chooses a partition $\Pi^{self}_{j^*}$ that bundles $v_{j^*} > v^{k^*}$ with values below v^{k^*}. Denote that bundle by $\pi^{self}_{j^*}$. 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 $\max_{j \neq j^*} v_j < \min_{v'_{j^*} \in \pi^{self}_{j^*}} v'_{j^*}$. 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.2.2 Proof of Theorem 1

To prove Theorem 1, we find necessary conditions that must be satisfied by an equilibrium information structure $\left(\tilde{\Pi}^{other}_{\lambda}, \tilde{\Pi}^{self}_{\lambda}\right)$. (Note that buyers can randomize over partitions, so the equilibrium information structure is not necessarily deterministic.) Lemmas 3 to 6 show, in a succession of steps, that an equilibrium can only put non-vanishing weight on information structures that have the form described in Theorem 1: $\Pi^{other}_{\lambda} \neq \{V\}$ and, for all $\pi^{other} \in \Pi^{other}_{\lambda}$,

$$(\star) \qquad \Pi^{self}_{\lambda}(\pi^{other}) = \{\{v_i \mid v_i < v^{\underline{k}}\}, \{v_i\}_{v^{\underline{k}} \leq v_i \leq v^{\overline{k}}}, \{v_i \mid v_i > v^{\overline{k}}\}\},$$

where $v^{\underline{k}} \equiv \min_{v_i' \in \pi^{other}} v_i'$ and $v^{\overline{k}} \equiv \max_{v_i' \in \pi^{other}} v_i'$. In words, after learning $\max_j v_j \in \pi^{other} \in \Pi_\lambda^{other}$, the agent bundles all the values he can have that are below (resp. above) his toughest competitor's valuation for sure. The following figure illustrates what such an information structure looks like.



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, in any equilibrium, if an information structure has probability $\Pr\left(\Pi_\lambda^{other}, \Pi_\lambda^{self}\right) \geq \varepsilon(\lambda)$, then it must have the following form: for all $\pi^{other} \in \Pi_\lambda^{other}$,*

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

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.[30]

Proof of Lemma 3. Fix λ, and let $\left(\Pi_\lambda^{other}, \Pi_\lambda^{self}\right)$ be an information partition that is chosen with probability at least ε in equilibrium. Take any $\pi^{other} \in \Pi_\lambda^{other}$, 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 $\Pi_\lambda^{self}\left(\pi^{other}\right)$ cannot bundle a value $v^* \in \pi^{other}$ with some other value weakly below $v^{\overline{k}}$. 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

[30]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.

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 $\underline{v} < \min_{v'_i \in \pi^{other}} v'_i \leq v^*$.

Step 1. We first show that, in any symmetric tremble-robust equilibrium, the agent's equilibrium bid at information partition $\left(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i\right)$ is bounded away from the boundary of the interval of valuations he deems possible: there exist $\eta', \eta'' > 0$ such that $\sigma_\lambda(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i) \in [\underline{v} + u_i + \eta', \overline{v} + u_i - \eta'']$ for almost all realizations of u_i.

Let $j^* \in \arg\max_j v_j$ be (any one of) agent i's toughest competitor(s).[31] 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. 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^*}. This implies that, for his equilibrium bid to be optimal, agent i must be indifferent between losing the auction and winning *at his equilibrium bid* (i.e., winning at a tie, such that he pays his equilibrium bid):

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

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.

We first argue that $\sigma_\lambda(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i) \geq \underline{v} + u_i + \eta'$ for some $\eta' > 0$. Indeed, at information set $\pi_i = (\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i)$ all values $v_i < \underline{v}$ have zero probability. Furthermore, value $v_i = v^* > \underline{v}$ has strictly positive probability since

$$\begin{aligned}&\Pr\left[v_i = v^* \mid \pi_i = \left(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i\right), \max_j \sigma(\pi_j) = \sigma_\lambda(\pi_i)\right] \\ &\quad = \frac{\Pr\left[\pi_i = (\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i), \max_j \sigma(\pi_j) = \sigma_\lambda(\pi_i) \mid v_i = v^*\right] \Pr(v_i = v^*)}{\Pr\left[\pi_i = (\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i), \max_j \sigma(\pi_j) = \sigma_\lambda(\pi_i)\right]}\end{aligned}$$

[31] 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.

$$= \frac{\Pr\left[\exists j^* : u_{j^*} = u_i\right] \Pr\left[\pi_i, \pi_{j^*} = \pi_i; \sigma(\pi_j) \leq \sigma_\lambda(\pi_i)\ \forall j \mid v_i = v^*, u_{j^*} = u_i\right] \Pr(v_i = v^*)}{\Pr\left[\exists j^* : u_{j^*} = u_i\right] \Pr\left[\pi_i = (\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i), \pi_{j^*} = \pi_i; \sigma(\pi_j) \leq \sigma_\lambda(\pi_i)\ \forall j \mid u_{j^*} = u_i\right]}$$

$$\geq \Pr\left[\pi_i, \pi_{j^*} = \pi_i; \sigma(\pi_j) \leq \sigma_\lambda(\pi_i)\ \forall j \mid v_i = v^*, u_{j^*} = u_i\right] \Pr(v_i = v^*)$$

$$\geq \Pr\left[\pi_i = \pi_j = (\pi^{other}, \pi^{self}_{\underline{v},\overline{v}})\ \forall j,\ u_j \leq u_i\ \forall j \neq j^* \mid v_i = v^*\right] \Pr(v_i = v^*)$$

$$\geq \varepsilon^{N-1} \left(\Pr(v_i = v^*)\right)^N \left(\Pr(u_j \leq u_i)\right)^{N-2},$$

which is strictly above zero for almost all u_i. Intuitively, when i wins *at his equilibrium bid*, he puts non-negligible weight on the state of the world in which all agents have the same valuation $v_j = v_i = v^*$, and the possibility that they all choose that same information structure we started with that has a least ε probability. Hence an agent's expected valuation at that information set, and then his bid, must be strictly above $\underline{v} + u_i$:

$$\sigma_\lambda(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i)$$

$$= \underline{v} + \sum_{\underline{v} < v_i \leq \overline{v}} (v_i - \underline{v}) \Pr\left[v_i \mid \pi_i = (\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}), \max_j \sigma(\pi_j) = \sigma_\lambda(\pi_i)\right] + u_i$$

$$\geq \underline{v} + (v^* - \underline{v})\varepsilon^{N-1} \left(\Pr(v_i = v^*)\right)^N + u_i.$$

Following a similar logic, we now argue that $\sigma_\lambda(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i) < \overline{v} + u_i - \eta''$ for some $\eta'' > 0$. This is the case if an agent with information set $\pi_i = (\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i)$ who knows that the highest bid among his competitors is $\sigma_\lambda(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i)$ still puts non-vanishing weight on the event $v_i = \underline{v}$:

$$\Pr\left[v_i = \underline{v} \mid \pi_i = (\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i), \max_j \sigma(\pi_j) = \sigma_\lambda(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i)\right]$$

$$\geq \frac{\Pr\left[\pi_i = (\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i), \max_j \sigma(\pi_j) = \sigma_\lambda(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i) \mid v_i = \underline{v}\right] \Pr(v_i = \underline{v})}{\Pr\left[\pi_i = (\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i), \max_j \sigma(\pi_j) = \sigma_\lambda(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i)\right]}$$

$$\geq \Pr\left(\Pi_j = \left(\Pi^{other}_\lambda, \Pi^{self}_\lambda\right) \forall j\right) \Pr\left[v_i = \underline{v}, \exists j', j'' \text{ s.t. } v_{j'} = v_{j''} = v^*, v_j = \underline{v} \forall j \neq j', j''\right]$$

$$\geq \varepsilon^{N-1} \left(\Pr(v_i = \underline{v})\right)^{N-3} \frac{(N-2)(N-1)}{2} \left(\Pr(v_i = v^*)\right)^2.$$

Hence $\sigma_\lambda(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i) \leq \overline{v} + u_i - (\overline{v} - \underline{v})\varepsilon^{N-1} \left(\Pr(v_i = \underline{v})\right)^{N-3} \frac{(N-2)(N-1)}{2} \left(\Pr(v_i = v^*)\right)^2$.

Step 2. 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. From step 1, we know the following: at information set $\pi_i = (\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_i)$, agent i knows that (i) with a strictly positive probability that is increasing in ε, the highest bid made by his competitors will be $\sigma_\lambda(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_j)$ for some u_j, and that (ii) that bid lies in $[\underline{v} + u_j + \eta', \overline{v} + u_j - \eta'']$, where both η' and η'' are increasing in ε and independent of λ. Hence there are strictly positive gains from unbundling values $v_i = \overline{v}$ and $v_i = \underline{v}$ whenever $(u_i - u_j) \in (-\eta'', \eta')$, as at the former buyer i wants to win at price $\sigma_\lambda(\pi^{other}, \pi^{self}_{\underline{v},\overline{v}}, u_j)$, whereas at the latter he does not. Since $\Pr[(u_i - u_j) \in (-\eta'', \eta')]$ is bounded away from zero, the expected gains from unbundling values $v_i = \overline{v}$ and $v_i = \underline{v}$ are as well.

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 between these values more than compensates the cost. Hence there exists $\varepsilon(\lambda)$ such that, if an information structure has probability $\Pr\left(\Pi^{other}_\lambda, \Pi^{self}_\lambda\right) \geq \varepsilon(\lambda)$, 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.

$\square$

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, in any equilibrium, if an information structure has probability $\Pr\left(\Pi^{other}_\lambda, \Pi^{self}_\lambda\right) \geq \varepsilon(\lambda)$, then it must have the following form:*

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

Proof of Lemma 4. Suppose not: there exists $\varepsilon > 0$ and a sequence of equilibria such that $\Pr\left(\Pi^{other}_\lambda, \Pi^{self}_\lambda\right) \geq \varepsilon$ for all λ and $\left\{\max_{v_i' \in \pi^{other}} v_i'\right\} \notin \Pi^{self}_\lambda \left(\pi^{other}\right)$ for some $\pi^{other} \in \Pi^{other}_\lambda$. 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 $\max_j v_j \in \pi^{other}$ then agent i knows that his toughest competitor has a valuation $\max_j v_j \in [\underline{v}, \overline{v}]$. Yet by assumption, agent i does not learn to distinguish $v_i = \overline{v}$.

We know from Lemma 4 that $\Pi^{self}_\lambda \left(\pi^{other}\right)$ cannot bundle $v_i = \overline{v}$ with lower values since an agent must learn to distinguish all the other values $v_i \in \pi^{other} \backslash \overline{v}$. Hence $\overline{v}$ must be bundles 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 must satisfy $\hat{v} > \overline{v}$ by the previous argument.

Step 1. We first show that, at information set $\pi_i = (\pi^{other}, \pi^{self}_{\geq \overline{v}}, u_i)$, agent i must bid arbitrarily close to $\overline{v} + u_i$ for sufficiently small λ. Note that, at this information set, 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. Hence agent i must be indifferent between losing and winning at a price equal to his equilibrium bid. Note however that if another agent j is at the same information set as i, then it must be that $v_i = \overline{v}$. Indeed, that other agent j must have learned that $\max_{k \neq j} v_k \in \pi^{other}$, which implies $v_i \leq \overline{v}$. When i ties to win, he then suffers a winner's curse and update his belief about his valuation to $v_i = \overline{v}$.

Step 2. We then show that such low 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 a agent i would bid (arbitrarily close to) $\overline{v} + u_i$ in all states of the world consistent with information set π_i. In particular, he would make such bids with probability at least ε 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 $v_i = \hat{v}$.

$\square$

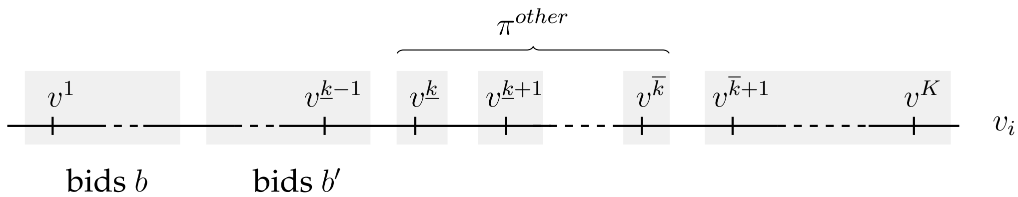
Lemma 5. *There exists $\overline{\lambda} > 0$ such that, for all $\lambda \leq \overline{\lambda}$, there exists $\varepsilon(\lambda) > 0$ with $\lim_{\lambda \longrightarrow 0} \varepsilon(\lambda) = 0$ such that, in any equilibrium, if an information structure has probability $\Pr\left(\Pi^{other}_\lambda,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

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

Proof of Lemma 5. Fix λ and let $\left(\Pi^{other}_\lambda, \Pi^{self}_\lambda\right)$ be an information partition that is chosen with probability at least ε in equilibrium. Take any $\pi^{other} \in \Pi^{other}_\lambda$, 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 $\Pr(b \leq \max_j \sigma_\lambda(\pi_j) \leq b' \mid \max_j v_j \in \pi^{other}, v_i < v^{\underline{k}}) > 0$, 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 $b' \leq v^{\underline{k}-1} + u_i < v^{\underline{k}} + \underline{u}$.[32] For j^* to make such a low bid in equilibrium, he must fail to learn that he has a high value and bundles his high value with lower ones. Let π_{j^*} the information set at which j^* acts in such a way, with $\sigma_\lambda(\pi_{j^*}) \leq b'$. 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,

[32] 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.

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.[33] 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: There exists $\pi^{self} \in \Pi^{self}(\pi^{other})$ such that $v_i \in \pi^{self}$ for all $v_i > v^{\overline{k}}$.

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 sometimes wins 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. □

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, in any symmetric equilibrium that is robust to trembles,*

$$\Pr\left(\Pi_{\lambda}^{other} = \{V\}\right) \leq \varepsilon(\lambda).$$

In words, an equilibrium can only put non-vanishing weight on information structures under which some information is acquired about others.

Proof of Lemma 6. Suppose not, and that the equilibrium puts non-vanishing probability on some information structure $\left(\Pi_{\lambda}^{other}, \Pi_{\lambda}^{self}\right)$ under which $\Pi_{\lambda}^{other} = \{V\}$. We know from Lemma 3 that buyers must then learn to distinguish all the values they might have

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

in π^{other}. Hence if buyers choose to acquire no information on others $\Pi_\lambda^{other} = \{V\}$, then they must become fully informed about their valuation $\Pi_\lambda^{self} = \{\{v\}_{v\in V}\}$.

We construct the same profitable deviation $\left(\widehat{\Pi}^{other}, \widehat{\Pi}^{self}, \widehat{\sigma}\right)$ as in the proof of Proposition 2. Consider the following information structure for buyer i, under which he first learns whether his toughest competitor has a value below or above some threshold v^*, and then only learns to distinguish values below v^* (resp, above v^*) if he learns that his toughest competitor has a value below v^* (resp, above v^*). We know that for some value of v^*, such information structure is strictly cheaper than becoming fully informed on self (Lemma 1). Hence deviating to that information structure leads to a strict reduction in information costs.

Doing so might however lead to a smaller gross payoff from the auction, if buyer i sometimes fails to learn his valuation for the good under this alternative information structure whereas he would learn it under $\left(\Pi_\lambda^{other}, \Pi_\lambda^{self}\right)$. We however show that this cannot be the case if $\Pr\left(\Pi_\lambda^{other}, \Pi_\lambda^{self}\right)$ is non-vanishing. There are two cases in which i fails to learn his valuation under the proposed deviation: either he learned his competitor has a value above the threshold $\max_j v_j > v^*$ while his lies below $v_i \leq v^*$, or the reverse. Let $j^* \in \arg\max_j v_j$.

Consider what happens in the first case. Under the proposed deviation, i only learns that $v_i \leq v^*$ whereas under Π_λ^{self} he learns his value fully and bids it. Bundling all this values can only lead to a different gross payoff for buyer i if the highest bid made by his competitors sometimes lies strictly below $v^* + \overline{u}$. In particular, this requires that buyer j^*, who has a value $v_{j^*} > v^* + \overline{u}$, sometimes makes a bid strictly below $v^* + \overline{u}$ and loses at that bid against i. However, recall that i chooses information structure $\left(\Pi_\lambda^{other}, \Pi_\lambda^{self}\right)$ with non-vanishing probability by assumption, in which case j^* loses from not learning his value and bidding it. If $\Pr\left(\Pi_\lambda^{other}, \Pi_\lambda^{self}\right)$ does not go to zero as λ goes to zero, then for λ small enough buyer j^* finds it profitable to learn his value so as to never lose against i in such a way. Hence if $\Pr\left(\Pi_\lambda^{other}, \Pi_\lambda^{self}\right)$ does not vanish, then for λ small enough, the proposed deviation leads to the exact same gross payoff for i in this first case.

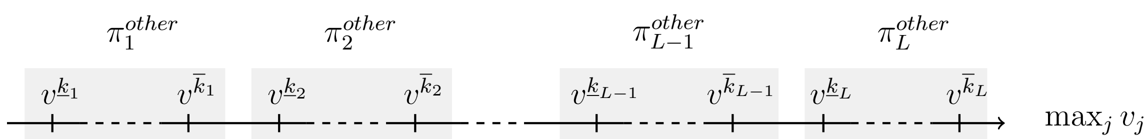
Now consider the second case, in which $\max_j v_j \leq v^*$ and $v_i > v^*$. Similarly as before, the alternative information structure can only lead to a lower gross payoff for i if he sometimes faces a bid strictly above $\min_{v_i > v^*} v_i + \underline{u}$ and loses at that bid. But if buyer i chooses the fully revealing partition with non-vanishing probability, that means that whoever is bidding high wins at a price strictly above his value with non-vanishing probability. Hence for λ small enough, that buyer would find it profitable to learn his valuation and avoid this loss, and so this cannot occur. There is again no loss in gross payoff from deviating to the strictly cheaper information structure, and the deviation we proposed must be strictly profitable.

□

B.2.3 Construction of a Symmetric Equilibrium

The previous subsection provides necessary conditions that must be satisfied by equilibrium information structures for small enough information cost λ. That however does not prove that an equilibrium with such structure exists. We now construct a symmetric tremble-robust equilibrium, thus proving existence.

Proposed Equilibrium Strategies. Let $\left(\Pi^{other}, \Pi^{self}\right)$ be a cost-minimizing information structure satisfying $(\star)$. Let $L = |\Pi^{other}|$ and denote the elements of the partition by $\Pi^{other} = \{\pi_l^{other}\}_{l=1}^{L}$. We know from Lemma 1 that $L > 1$. Denote by $v^{\underline{k}_l}$ and $v^{\overline{k}_l}$ the lowest and highest values in π_l^{other}, respectively, and order the elements of the partition in "increasing order," meaning that $\overline{k}_{l-1} + 1 = \underline{k}_l$:



Let $\overline{l} \equiv \min_l \left\{ l \,\middle|\, v^{\overline{k}_l} > \mathbb{E}\left[v_i \mid v_i < v^{\underline{k}_{l'}}\right] \forall l' \right\}$.[34]

To gain intuition about the equilibrium construction, we first argue why buyers cannot *always* choose $\Pi^{self}(\pi_l^{other}) = \left\{ \{v_i : v_i < v^{\underline{k}_l}\}, \{v_i\}_{v^{\underline{k}_l} \leq v_i \leq v^{\overline{k}_l}}, \{v_i : v_i > v^{\overline{k}_l}\} \right\}$ after learning $\max_j v_j \in \pi_l^{other}$. Suppose they do, and consider what happens when i learns $\max_j v_j \in \pi_l^{other}$ for some $l \geq \overline{l}$. For instance, $\max_j v_j \in \pi_{L-1}^{other}$. According to $\Pi^{self}(\pi_l^{other})$, buyer i should learn to distinguish $\{v^{\overline{k}_l}\}$ from $\{v_i : v_i > v^{\overline{k}_l}\}$. However, he knows that the highest bid he might face from his competitors is $v^{\overline{k}_l} + u_j$, and that he can only face such high bid if that competitor j learns his value, which is only the case if j learned that $\max_{j' \neq j} v_j \in \pi_l^{other}$. Since i is j's toughest competitor, that can only occur when i's

[34] It can be that $\overline{l} = 1$ (this is for instance the case if $L = 2$) but it can be strictly greater.

value is $v_i = v^{\bar{k}_l}$. Buyer i then achieves the same gross payoff by bundling all $v_i \geq v^{\bar{k}_l}$, and bidding $v^{\bar{k}_l} + u_i$. Since this leads to a strict reduction in information costs, this is a profitable deviation. To avoid such deviation, there must be some (vanishing) mixing in equilibrium.

Consider a symmetric strategy profile in which all buyers choose with probability one $\Pi^{other} = \{\pi_l^{other}\}_{l=1}^{L}$. After learning $\max_j v_j \in \pi_l^{other}$ with $l < \bar{l}$, each buyer chooses

$$\Pi^{self}(\pi_l^{other}) = \left\{\{v_i : v_i < v^{\underline{k}_l}\}, \{v_i\}_{v^{\underline{k}_l} \leq v_i \leq v^{\bar{k}_l}}, \{v_i : v_i > v^{\bar{k}_l}\}\right\},$$

with probability one. After learning $\max_j v_j \in \pi_l^{other}$ with $l \geq \bar{l}$, buyers randomize over the following information partitions about self:

- With (vanishing) probability $q_{\geq l}$, they choose

$$\Pi_{\geq l}^{self}(\pi_l^{other}) \equiv \left\{\{v_i : v_i < v^{\underline{k}_l}\}, \{v_i\}_{v^{\underline{k}_l} \leq v_i < v^{\bar{k}_l}}, \{v_i : v_i \geq v^{\bar{k}_l}\}\right\} \quad \text{if } \bar{l} \leq l < L.$$

- With (vanishing) probability $q_{<l}$, they choose

$$\Pi_{<l}^{self}(\pi_l^{other}) \equiv \left\{\{v_i : v_i < v^{\bar{k}_{l-1}}\}, \{v^{\bar{k}_{l-1}}\}, \{v_i\}_{v^{\underline{k}_l} \leq v_i \leq v^{\bar{k}_l}}, \{v_i : v_i > v^{\bar{k}_l}\}\right\} \quad \text{if } \bar{l} < l \leq L.$$

- With remaining (close to one) probability, they choose

$$\Pi^{self}(\pi_l^{other}) \equiv \left\{\{v_i : v_i < v^{\underline{k}_l}\}, \{v_i\}_{v^{\underline{k}_l} \leq v_i \leq v^{\bar{k}_l}}, \{v_i : v_i > v^{\bar{k}_l}\}\right\}.$$

At almost all information sets $\left(\pi^{other}, \pi^{self}, u_i\right)$, buyers simply bid their expected valuation for the good given the information they acquired:

$$\sigma\left(\pi^{other}, \pi^{self}, u_i\right) = \mathbb{E}\left[v_i \mid v_i \in \pi^{self}\right] + u_i.$$

The only exception is when $\pi^{self} = \{v_i : v_i \geq v^{\bar{k}_l}\}$, in which case buyers bid

$$\begin{aligned}
\sigma(\pi_i) &= \mathbb{E}[v_i \mid \pi_i, \max_j \sigma(\pi_j) = \sigma(\pi_i)] + u_i \\
&= v^{\bar{k}_l} + u_i + \underbrace{\left(\mathbb{E}(v_i \mid v_i \in \pi_{l+1}^{other}) - v^{\bar{k}_l}\right) \Pr\left(v_i \in \pi_{l+1}^{other} \,\middle|\, \pi_i, \max_j \sigma(\pi_j) = \sigma(\pi_i)\right)}_{\equiv WB(u_i, q_{<l+1})}
\end{aligned}$$

where $WB(u_i, q_{<l+1})$ is the winner's blessing that i faces upon tying, which goes to zero when $q_{<l+1}$ goes to zero. Indeed, the winner's blessing solves

$$
\begin{aligned}
WB(u_i, q_{<l+1}) &= \left(\mathbb{E}(v_i \mid v_i \in \pi^{other}_{l+1}) - v^{\bar{k}_l}\right) \\
&\times \frac{\Pr\left(\max_{\substack{j: v_j = v^{\bar{k}_l}, \\ \Pi^{self}_j = \Pi^{self}_{<l+1}}} u_j = u_i + WB(u_i, q_{<l+1})\right) \Pr\left(\max_j v_j = v^{\bar{k}_l}\right) \Pr\left(v_i \in \pi^{other}_{l+1}\right)}{\Pr\left[\pi_i, \max_j \sigma(\pi_j) = \sigma(\pi_i)\right]}.
\end{aligned}
$$

Note that the numerator goes to zero as $q_{<l+1}$ go to zero, as it requires some of i's competitors to have chosen partition $\Pi^{self}_{<l+1}$. On the contrary, the denominator does not scale with $q_{<l+1}$.

Checking Consistency of Proposed Strategies. To check that this is indeed an equilibrium, we first need to prove that at every possible information set, the bidding strategy $\sigma\left(\pi^{other}, \pi^{self}, u_i\right)$ is optimal given that others follow the equilibrium strategy. This is direct when a buyer fully learns his valuation, i.e. when $|\pi^{self}| = 1$, by the strategy-proofness of a second-price auction.

When buyers fail to fully learn their valuations $|\pi^{self}| > 1$, it can in principle be possible that others' bids provide them with additional information about their own value, leading to a winner's curse/winner's blessing. In equilibrium, this only happens when $\pi^{self} = \{v_i : v_i \geq v^{\bar{k}_l}\}$ for some $l \geq \bar{l}$, and we leave that case for the end. There are only two other possible cases for $|\pi^{self}| > 1$:

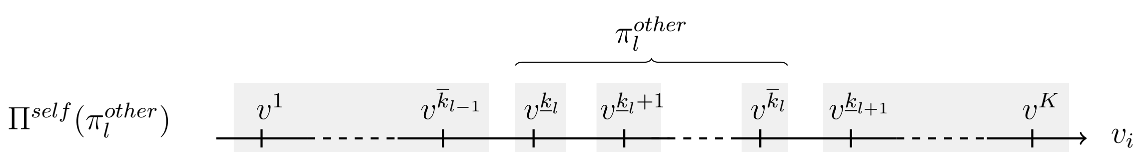
- If $\pi^{self} = \{v_i : v_i < v^{\underline{k}_l}\}$, or $\pi^{self} = \{v_i : v_i < v^{\bar{k}_{l-1}}\}$ for $\bar{l} < l$, the equilibrium requires that buyer i bids $\mathbb{E}[v_i \mid \pi^{self}] + u_i$. We show that i loses with probability one at that bid in equilibrium. Note that i can only have such information set if one of i's competitors, call him j^*, has a strictly greater value than i. Given the equilibrium information structure, buyer j^* can only have the following information sets: either $\pi^{self}_{j^*} = \{v_{j^*}\}$ or $\pi^{self}_{j^*} = \{v_{j^*} : v_{j^*} > v^{\bar{k}_l}\}$, or $\pi^{self}_{j^*} = \{v_{j^*} : v_{j^*} \geq v^{\bar{k}_l}\}$. Under all three information sets, buyer j^* makes a bid strictly above i's bid by definition of $\bar{l}$. Since i loses with probability one, our equilibrium refinement requires that i's bid be optimal against vanishing trembles that have full support and *that are independent of agents' information sets*—hence winning against trembled-bids reveals no information to i about his value, and it is strictly optimal for i to bid his expected valuation for the good given his information set.

- If $\pi^{self} = \{v_i : v_i > v^{\bar{k}_l}\}$, the equilibrium requires that buyer i bids $\mathbb{E}\left[v_i \mid v_i > v^{\bar{k}_l}\right] + u_i$. Buyer i must win with probability one at that bid in equilibrium. Indeed, all others buyers can only be at one of the following information sets: $\pi_j^{self} = \{v^{\bar{k}_{l'}}\}$ for some $l' \leq l$, or $\pi_j^{self} = \{v_j : v_j < v^{\underline{k}_{l'}}\}$. At either of them, they make a bid strictly below i's. Similarly as above, i's equilibrium bid must be optimal in case (at least) one of i's competitors trembles. Since trembles have full support, it is strictly optimal for i to bid his expected valuation given his information set π_i.

There is only one information set at which buyers' equilibrium bid reflects a winner's curse or blessing: $\pi^{self} = \{v_i : v_i \geq v^{\bar{k}_l}\}$ for $l \geq \bar{l}$.[35] Buyers then bid their expected valuations conditional on winning *at their equilibrium bid*. As explained above, such bid is arbitrarily close to $v^{\bar{k}_l} + u_i$ for $q_{<l+1}$ small enough. In particular, for every $u_i < \overline{u}$, there exists $q_{<l+1}$ small enough such that i's expected valuation conditional on tying as calculated above lies in $[v^{\bar{k}_l} + \underline{u}, v^{\bar{k}_l} + \overline{u}]$. Because they are indifferent between winning and losing at such bid, they have no incentive to marginally change their bid. They also never face a bid above this interval, and so deviating to an even higher bid cannot make them better off.

Now we need to check that, given that others follow the equilibrium strategy, the information structure that buyers choose is optimal. Take any $\pi_l^{other} \in \Pi^{other}$. We check that the proposed (randomization over) information partition about self is optimal.

First consider the case where $l < \bar{l}$. The proposed equilibrium puts weight on a single partition about self:

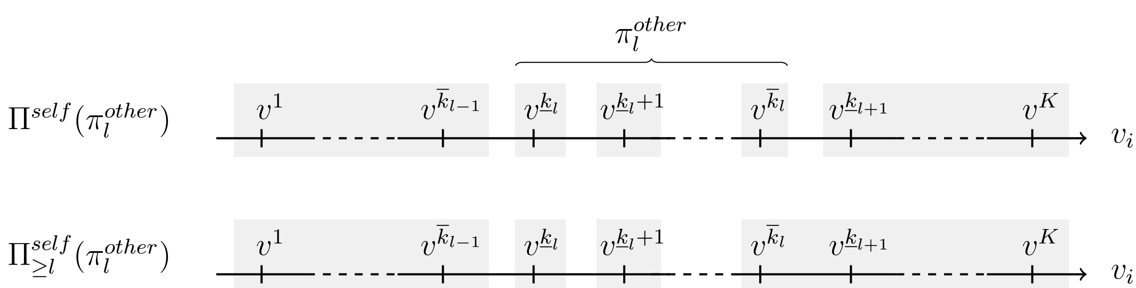


We show that no other partition can make a buyer i strictly better off. First, any partition that bundles two valuations $v_i, v_i' \in \pi_l^{other}$ must make i worse off, as at such information set his toughest competitor would learn his valuation and bid in that range.

[35] The equilibrium probability of this information set vanishes as λ goes to zero.

Second, i's toughest competitor's bid must be weakly above $v^{\underline{k}_l} + \underline{u}$, and so learning to distinguish some values $v_i < v^{\underline{k}_l}$ strictly increases information costs without leading to any strict gains in gross payoff. Things are a bit more subtle if i has a valuation above $v^{\overline{k}_l}$. He might then face a bid above $v^{\overline{k}_l} + \overline{u}$, but, importantly, would always want to win at that bid. Indeed, this only happens when others fail to learn their valuations, but only learn that their valuations lie below i's: that is, if $v_i \in \pi_{l'}^{other}$ for l' large enough and $\pi_j^{self} = \{v_j : v_j < v^{\underline{k}_{l'}}\}$ for some j, such that j bids $\mathbb{E}[v_j \mid v_j < v^{\underline{k}_{l'}}] + u_j > v^{\overline{k}_l} + \overline{u}$. Since $\sigma(\pi_l^{other}, \{v_i : v_i > v^{\overline{k}_l}\}, u_i) = \mathbb{E}[v_i \mid v_i > v^{\overline{k}_l}] + u_i$, i must be winning the auction with probability one at this information, and so learning to distinguish some values $v_i > v^{\overline{k}_l}$ cannot lead to strict gains in gross payoff. The only deviation from $\Pi^{self}(\pi_l^{other})$ to consider is that of bundling $\{v^{\overline{k}_l}\}$ with $\{v_i : v_i > v^{\overline{k}_l}\}$. Doing so leads to a strict reduction in information costs. It however also entails a strict, non-vanishing loss in gross payoffs. Indeed, if i were to make such bundle, he would either (i) bid sufficiently high so as to always win, in which case he would also win the auction whenever $\nu_i = v^{\overline{k}_l} + u_i < \max_j \nu_j = v^{\overline{k}_l} + u_j$ and incur a strict loss, or (ii) bid sufficiently low so as to not incur that loss (i.e., at $v^{\overline{k}_l} + u_i$) but then lose the auction with strictly positive probability when some j bids $\mathbb{E}[v_j \mid v_j < v^{\underline{k}_{l'}}] + u_j > v^{\overline{k}_l} + \overline{u}$.[36] For lambda small enough, the loss in gross payoffs must be greater than the decrease in information costs, and such deviation cannot be profitable.

Now consider the case where $l \geq \overline{l}$. Our proposed equilibrium requires buyers to randomize over several Π^{self}. In particular, a buyer must be indifferent between distinguishing value $\{v^{\overline{k}_l}\}$ from $\{v_i : v_i > v^{\overline{k}_l}\}$, and bundling all these values together:



If he bundles these values, he makes a bid that is arbitrarily close to $v^{\overline{k}_l} + u_i$ for sufficiently small $q_{<l+1}$. The opportunity cost of such bundle is that he might lose if one of his competitors has a value above i's bid $v^{\overline{k}_l} + u_j > v^{\overline{k}_l} + u_i$, while his value might

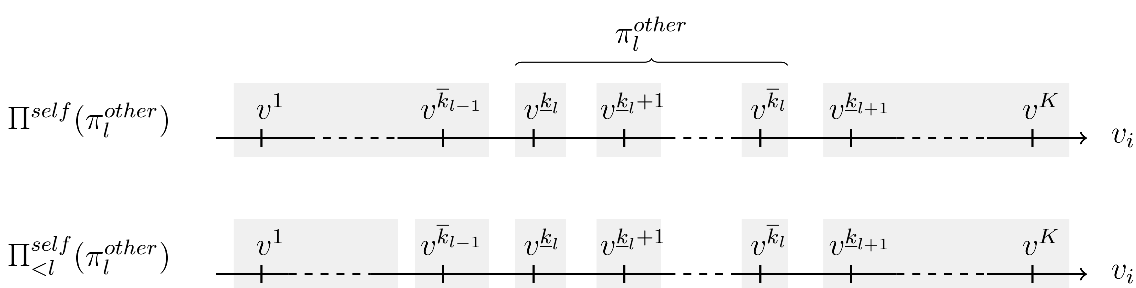
[36] This must occur in equilibrium since $l < \overline{l}$.

have been (with vanishing probability) strictly greater. He is indifferent between the two if:

$$
\begin{aligned}
&\Pr\left(v_i \in \pi_{l+1}^{other}\right) \Pr\left[\max_j \sigma(\pi_j) > \sigma\left(\pi_l^{other}, \{v_i \mid v_i \geq v^{\bar{k}_l}\}, u_i\right) \;\middle|\; \max_j v_j \in \pi_l^{other}, v_i \in \pi_{l+1}^{other}\right] \\
&\qquad \times \mathbb{E}\left[v_i + u_i - \max_j \sigma(\pi_j) \;\middle|\; v_i \in \pi_{l+1}^{other}, \max_j \sigma(\pi_j) > \sigma\left(\pi_l^{other}, \{v_i \mid v_i \geq v^{\bar{k}_l}\}, u_i\right)\right] \\
&+ \Pr\left(v_i = v^{\bar{k}_l}\right) \Pr\left[v^{\bar{k}_l} + u_i < \max_j \sigma(\pi_j) \leq \sigma\left(\pi_l^{other}, \{v_i \mid v_i \geq v^{\bar{k}_l}\}, u_i\right) \mid \max_j v_j \in \pi_l^{other}\right] \\
&\qquad \times \mathbb{E}\left[v^{\bar{k}_l} + u_i - \max_j \sigma(\pi_j) \;\middle|\; v^{\bar{k}_l} + u_i < \max_j \sigma(\pi_j) \leq \sigma\left(\pi_l^{other}, \{v_i \mid v_i \geq v^{\bar{k}_l}\}, u_i\right)\right] \\
&= \lambda\left[c\left(\Pi^{self}(\pi_l^{other}), p\right) - c\left(\Pi_{\geq l}^{self}(\pi_l^{other}), p\right)\right] > 0.
\end{aligned}
$$

Recall that $\sigma\left(\pi_l^{other}, \{v_i \mid v_i \geq v^{\bar{k}_l}\}, u_i\right)$ goes to $v^{\bar{k}_l} + u_i$ as $q_{<l+1}$ goes to zero. Furthermore, since other bidders lose with probability one when $v_i > v^{\bar{k}_l}$ under all information partitions that have non-vanishing weight, the LHS is strictly decreasing in $q_{<l+1}$ and goes to zero with $q_{<l+1}$. For lambda sufficiently small, there thus exists $q_{<l+1} \in (0,1)$ such that the equation holds, and $q_{<l+1}$ goes to zero as λ goes to zero.

Finally, a buyer must be indifferent between distinguishing value $v_i = v^{\bar{k}_{l-1}}$ from $\{v_i : v_i < v^{\bar{k}_{l-1}}\}$, or not:



The gain from learning to distinguish these values is that i's toughest competitor j^* might (with vanishing probability) choose $\Pi_{>l-1}^{self}$ and bid arbitrarily close to $v^{\bar{k}_{l-1}} + u_{j^*}$. There is then potential value for buyer i to match that bid when $v_i = v^{\bar{k}_{l-1}}$. A buyer is then indifferent if and only if

$$
\Pr\left(v_i = v^{\bar{k}_{l-1}}\right) \Pr\left(\max_j \sigma(\pi_j) \leq v^{\bar{k}_{l-1}} + u_i \;\middle|\; \max_j v_j \in \pi_l^{other}\right)
$$

$$\times \mathbb{E}\left[v^{\bar{k}_{l-1}} + u_i - \max_j \sigma(\pi_j) \,\middle|\, \max_j \sigma(\pi_j) \le v^{\bar{k}_{l-1}} + u_i, \max_j v_j \in \pi_l^{other} \right]$$

$$= \lambda \left[c\left(\Pi_{<l}^{self}(\pi_l^{other}), p \right) - c\left(\Pi^{self}(\pi_l^{other}), p \right) \right] > 0.$$

Note that $\Pr\left(\max_j \sigma(\pi_j) \le v^{\bar{k}_{l-1}} + u_i \,\middle|\, \max_j v_j \in \pi_l^{other} \right)$ goes to zero as $q_{\ge l-1}$ goes to zero, as buyer j^* only makes a bid that might be below $v^{\bar{k}_{l-1}} + u_i$ when he chooses information partition $\Pi_{\ge l-1}^{self}$. For lambda sufficiently small, there thus exists $q_{\ge l-1} \in (0,1)$ such that the equation holds, and $q_{\ge l-1}$ goes to zero as λ goes to zero.

Finally, we argue that buyers do not want to deviate to another information partition about others. By construction, their prescribed equilibrium partition Π^{other} is one that leads to the smallest overall information costs. All information that has strictly positive value is furthermore acquired by previous arguments, and so buyers cannot find it profitable to deviate to another partition.

B.3 Proof of Results of Section 4

B.3.1 Proof of Theorem 2

Lemma 7. *There exists $\bar{\lambda} > 0$ such that, for all $\lambda \le \bar{\lambda}$, there exists $\varepsilon(\lambda) > 0$ with $\lim_{\lambda \longrightarrow 0} \varepsilon(\lambda) = 0$ such that, in any symmetric equilibrium that is robust to trembles,*

$$\Pr\left(\Pi_\lambda^{other} = \{\{v^1\}, \{v^2, \ldots, v^K\}\} \right) \le \varepsilon(\lambda).$$

Proof of Lemma 7. The proof of Lemma 7 is very similar to that of Lemma 6. Suppose not, and that the equilibrium puts non-vanishing probability on some information structure $\left(\Pi_\lambda^{other}, \Pi_\lambda^{self} \right)$ under which $\Pi_\lambda^{other} = \{\{v^1\}, \{v^2, \ldots, v^K\}\}$. Buyers must then become fully informed about their own valuation if they learn that their toughest competitor does not have the lowest possible valuation $\Pi_\lambda^{self}(\{v^2, \ldots, v^K\}) = \{\{v\}_{v \in V}\}$, and choose $\Pi_\lambda^{self}(\{v^1\}) = \{\{v^1\}, \{v^2, \ldots, v^K\}\}$ otherwise.

As before, we construct a profitable deviation $\left(\widehat{\Pi}^{other}, \widehat{\Pi}^{self}, \widehat{\sigma} \right)$ in which $\widehat{\Pi}^{other} = \{\{v^1\}, \{v^2, \ldots, v^*\}, \{v : v > v^*\}\}$ for some $v^* \in (v^1, v^K)$, and $\left(\widehat{\Pi}^{other}, \widehat{\Pi}^{self} \right)$ satisfies $(\star)$. We know from Lemma 2 that for some v^*, such information structure is strictly cheaper than the one we started with $\left(\Pi_\lambda^{other}, \Pi_\lambda^{self} \right)$.

We furthermore show that deviating to this cheaper information structure cannot strictly reduce i's gross payoff. If $\max_j v_j = v^1$, then i chooses the very same information structure about self under $\widehat{\Pi}^{self}$ and Π_λ^{self}, and so the two must yield the same gross payoff. He can only be made worse if he fails to learn his value under $\widehat{\Pi}^{self}$ while he learns it under Π_λ^{self}. This is the case if $\max_j v_j > v^*$ but $v_i \leq v^*$, or if $\max_j v_j \in [v^2, v^*]$ but $v_i > v^*$. However, we can use the very same argument as in the proof of Lemma 6 to show that these bundles cannot make i strictly worse off in equilibrium, and hence that deviating to the strictly cheaper information structure strictly increases i's overall payoff. $\square$

Proof of Theorem 2. We know from Lemmas 3 to 6 that, for λ small enough, the only information structures that have non-trivial probability must satisfy $\Pi_\lambda^{other} \neq \{V\}$ and, for all $\pi^{other} \in \Pi_\lambda^{other}$,

$$(\star) \qquad \Pi_\lambda^{self}(\pi^{other}) = \{\{v_i \mid v_i < v^{\underline{k}}\}, \{v_i\}_{v^{\underline{k}} \leq v_i \leq v^{\bar{k}}}, \{v_i \mid v_i > v^{\bar{k}}\}\},$$

where $v^{\underline{k}} \equiv \min_{v_i' \in \pi^{other}} v_i'$ and $v^{\bar{k}} \equiv \max_{v_i' \in \pi^{other}} v_i'$. That is, they involve acquiring some information about others, and then only differentiating between own valuations that fall in the same range as the toughest competitor's valuation. 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 buyers fail to learn their valuations fully but only learn that they do 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 they bid their expected valuations given their information sets in any equilibrium that is robust to trembles. Denote that bundle by $\pi_<^{self}$. We want to show that, at information set $\pi_i = (\pi^{other}, \pi_<^{self}, u_i)$, buyers bid $\mathbb{E}[v_i \mid v_i < \min_{v_i' \in \pi^{other}} v_i'] + u_i$. We already know that, in any tremble-robust equilibrium, buyers cannot make a bid that lies outside the interval of values they deem possible:

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

Furthermore, it has to be that a buyer i with that information set loses 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 $\sigma_\lambda(\pi^{other}, \pi^{self}_<, u_i) < \min_{v_i' \in \pi^{other}} v_i' + \underline{u} \leq \nu_{j^*}$. 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 might tremble 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 true expected valuation for the good given his information set:

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

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.[37] Take any symmetric tremble-robust equilibrium, and denote by $q(\lambda)$ the probability that a buyer chooses an information structure satisfying ($\star$). We know from Lemmas 3 to 6 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 probability, and induces a revenue that is bounded above by the highest possible valuation $\max_{v \in V} v + \overline{u}$. 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 the noise terms

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

u_i's is continuous.) Similarly, let i_2 the second-highest valuation buyer. 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 $b_{(2)} = \nu_{j^*} \leq \mathbb{E}[\nu_{(2)} \mid \nu_{j^*}, \nu_{j^*} \leq \nu_{i_1}]$. In the second case, the second-highest bidder failed to learn his value, which means that $\pi_{j^*}^{self} = \{v_j \mid v_j \leq v^*\}$ for some v^*. We know from Step 1 that j^* must have bid $\sigma_\lambda(\pi_{j^*}) = \mathbb{E}[\nu_{j^*} \mid v_{j^*} \leq v^*]$, which always lies weakly below $\mathbb{E}[\nu_{(2)} \mid v_{j^*} \leq v^*, \nu_{j^*} \leq \nu_{i_1}]$.

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. Take any information structure $(\Pi_\lambda^{other}, \Pi_\lambda^{self})$ that has non-vanishing weight in equilibrium. Let $\underline{v} = \min\{v \mid \exists \pi^{other} \in \Pi_\lambda^{other} \text{ s.t. } v \in \pi^{other}, v^K \in \pi^{other}\}$ denote the smallest valuation that Π_λ^{other} bundles with v^K. We know from Lemmas 6 and 7 that $\Pi_\lambda^{other} \neq \{V\}$ and $\Pi_\lambda^{other} \neq \{\{v^1\}, \{v^2, \ldots, v^K\}\}$. Hence $\underline{v} > v^2$. Consider what happens when $v_{i_1} \geq \underline{v}$ and $v_j < \underline{v}$ for all $j \neq i_1$. When all buyers choose this information structure, all buyers $i \neq i_1$ 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: $|\pi_j^{self}| \geq 2$ as $\pi_j^{self} = \{v_j \mid v_j < \underline{v}\}$. The second-highest bid then equals $\mathbb{E}[v_j \mid v_j < \underline{v}] + \max_{j \neq i_1} u_j$.

Overall, we get

$$\mathbb{E}\left[\text{ equilibrium revenue } \mid \left(\Pi_\lambda^{other}, \Pi_\lambda^{self}\right)\right] - \mathbb{E}[\nu_{(2)}] \leq \left(1 - [q(\lambda)]^N\right)\left(v^K - \mathbb{E}[\nu_{(2)}]\right)$$

$$+ [q(\lambda)]^N \Pr\left(\Pi_\lambda^{other}, \Pi_\lambda^{self}\right)^N \times$$

$$\Pr\left(v_{(1)} \geq \underline{v}, v_{(2)} < \underline{v}\right)\left(\mathbb{E}[v_i \mid v_i < \underline{v}] - \mathbb{E}[v_{(2)} \mid v_{(2)} < \underline{v}, v_{(1)} \geq \underline{v}] + \overline{u} - \underline{u}\right).$$

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

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

Then

$$\mathbb{E}\left[\text{ equilibrium revenue} \mid \left(\Pi_\lambda^{other}, \Pi_\lambda^{self}\right)\right] - \mathbb{E}[\nu_{(2)}] \leq \left(1 - [q(\lambda)]^N\right)\left(v^K - \mathbb{E}[\nu_{(2)}]\right)$$

$$+ [q(\lambda)]^N \Pr\left(\Pi_\lambda^{other}, \Pi_\lambda^{self}\right)^N \Pr\left(v_{(1)} \geq \underline{v}, v_{(2)} < \underline{v}\right) \Pr\left(v_{(1)} \geq \underline{v}, v_{(2)} < \underline{v}\right)\left(-l + \overline{u} - \underline{u}\right).$$

For $\overline{u}$, $\underline{u}$ small enough, $(-l + \overline{u} - \underline{u}) < 0$. Since $\lim_{\lambda \longrightarrow 0} q(\lambda) = 1$, the first RHS term goes to zero as information costs vanish. Since $\lim_{\lambda \longrightarrow 0} \Pr\left(\Pi_\lambda^{other}, \Pi_\lambda^{self}\right) > 0$, 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 Results of Section 4.2

Proof of Proposition 3. Take any symmetric equilibrium that satisfies our equilibrium refinement $\left(\tilde{\Pi}_\lambda^{other}, \tilde{\Pi}_\lambda^{self}, \sigma_\lambda\right)$, and fix some realization of $(v_i)_i$. We show that two bidders i_1, i_2 with values $v_{i_1} < v_{i_2}$ cannot both enter the auction with probability greater than $\varepsilon > 0$ for all λ. Given a couple of information partitions $\left(\Pi^{other}, \Pi^{self}\right)$, let $\pi(v) \in \Pi^{other} \times \Pi^{self}$ denote the information set at which a buyer must be in state $v = (v_i)_i$.[38] Formally, we show that

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

cannot be greater than ε irrespective of λ for both i_1 and i_2 if $v_{i_1} < v_{i_2}$. 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 $\sigma_\lambda(\pi_i) > 0$ for $i = i_1, i_2$.

Step 1. We first show that i_1 must know and bid his valuation at this information set: $\pi_{i_1}^{self} = \{v_{i_1}\}$ and $\sigma_\lambda(\pi_{i_1}) = v_{i_1} + u_{i_1}$. Suppose not: at this information set i_1 bundles some valuations he might have $|\pi_{i_1}^{self}| > 1$, and let $\underline{v}$ and $\overline{v}$ denote the lowest and

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

highest values in π_{i_1}, respectively. For i_1 to find it optimal to enter the auction and bid $\sigma_\lambda(\pi_{i_1}) > 0$, it must be that he wins with strictly positive probability at such bid. In particular, this implies that i_2 is sometimes at an information set at which he bids lower than i_1. At information set π_{i_1}, buyer i_1 cannot rule out the possibility that some other buyer has the same value as him: $v_j = v_{i_1}$ for some j. Indeed, $\pi_{i_1}^{other}$ is only informative of $\max_{j\neq i_1} v_j \geq v_{i_2} > v_{i_1}$. Hence with non-vanishing probability, one of i_1's opponents has the same information set as him $\pi_j = \pi_{i_1}$ and makes the same equilibrium bid. With non-vanishing probability, buyer i_1 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 $|\pi_{i_1}^{self}| > 1$.

Step 2. We now show that in state $v = (v_i)_i$, i_2 must be entering the auction and outbidding i_1 with probability one. Take any information set $\pi(v_{i_2}, v_{-i_2})$ that i_2 might have in equilibrium given the realized $(v_i)_i$. By assumption, one of these information sets is the one we started with π_{i_2} at which he enters. At this information set, he must be outbidding i_1. Indeed we know that i_1 is bidding $v_{i_1} + u_{i_1} < \nu_{i_2}$, and that i_1 wins sufficiently often at that bid to justify its entry. So there are strictly positive, non-vanishing gains from outbidding i_1, and for λ small enough, i_2 must learn sufficiently about his valuation to do so.

Suppose that at some other information set $\pi(v_{i_2}, v_{-i_2})$ that has positive probability in equilibrium, buyer i_2 stays out of the auction in that state of the world, and gets zero gross payoff. By choosing a finer information structure, buyer i_2 could have entered and gotten a strictly positive, non-vanishing gross payoff. Indeed, since i_2 has a strictly greater value than i_1, i_2 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.

We now argue that i_1 cannot find it profitable to enter the auction at information set π_{i_1}. There are two cases: either $\pi_{i_1}^{other}$ is sufficiently fine that i_1 can predict that i_2 will enter with probability one and bid higher than ν_{i_1}, or not. In the first case, i_1 cannot find it optimal to pay the entry fee κ. In the second, i_1 could deviate to a finer $\pi_{i_1}^{other}$ 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 $\pi_{i_1}^{other}$, and this deviation leads to a strictly higher overall payoff. $\square$

B.4 Proof of Results of Section 5

Proof of Theorem 3. Let $\Pi_\lambda^{other} = \{\{\pi_l^{other}\}_{l=1}^L\}$ 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 3 has two steps. We first show that, for N large enough, buyers learn to distinguish whether or not their toughest competitor has a value equal to v^K (i.e., $\{v^K\} \in \Pi^{other}$) as λ goes to zero. 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: *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 every N, $\{v^K\} \notin \Pi_\lambda^{other}$ even as λ goes to zero. In words, the equilibrium partition that buyers choose about others bundles v^K with some other possible values: $\{v^{\underline{k}_L}, \dots, v^K\} \in \Pi_\lambda^{other}$ for some $v^{\underline{k}_L} < v^K$.

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

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

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

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

Consider an alternative information partition about others that unbundles v^K from

$\{v^{\underline{k}_L}, \ldots, v^{K-1}\}$, but keeps the rest of the partition the same:

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

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

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

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 $\max_j v_j \in \{v^{\underline{k}_L}, \ldots, v^K\}$ 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, $\lim_{N\longrightarrow\infty} \delta_{\{v^{\underline{k}_L},\ldots,v^K\}} = \delta_{\{v^K\}}$. 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.[39]

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 $\max_j v_j \in \{v^{\underline{k}_L}, \ldots, v^K\}$, then we know from Theorem 1 that, for λ sufficiently small, he must choose to partition his own set of valuations into $\{\{v^1, \ldots, v^{\bar{k}_{L-1}}\}, \{v^k\}_{k=\underline{k}_L,\ldots,K}\}$. That is, he bundles all the values that he knows are below the toughest opponent's, and learns to distinguish all the values he can have in the interval $[v^{\underline{k}_L}, v^K]$. If the buyer learns that $\max_j v_j \in \{v^{\underline{k}_L}, \ldots, v^{K-1}\}$, 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: $\{\{v^1, \ldots, v^{K-1}\}, \{v^K\}\}$. Indeed, such partition cannot make him worse off than $\{\{v^1, \ldots, v^{\bar{k}_{L-1}}\}, \{v^k\}_{k=\underline{k}_L,\ldots,K}\}$.[40] Hence the

[39]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.

[40]For him to be worse off, it would have to be that the highest bid he faces sometimes lies strictly below $v^K + \underline{u}$, and that by choosing $\{\{v^1, \ldots, v^{\bar{k}_{L-1}}\}, \{v^k\}_{k=\underline{k}_L,\ldots,K}\}$ he sometimes wins at that bid. In particular, that requires i's toughest competitor j^*, whom we know has value $v_{j^*} = v^K$ to sometimes fail to learn his value and make a low bid. However, since the information structure we started with is chosen with non-vanishing probability, that means j^* loses against i's bid with non-vanishing probability.

gain in information cost on self is

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

Since $\Pr(\max_j v_j = v^K)$ tends to one as N tends to infinity, this tends to the strictly positive expression in parenthesis. For N large enough, the gains associated with cheaper information acquisition on self more than compensate the cost of the extra information about others, and $\widehat{\Pi}^{other}$ yields a strictly higher payoff to the buyer than Π^{other}_λ.

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 $r = v^K - \varepsilon$ for ε small ($|\underline{u}| < \varepsilon < v^K - v^{K-1}$). 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

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

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 $r = v^K - \varepsilon$ then yields an expected revenue of

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

Now consider what happens in equilibrium if no reserve price is imposed, but there are $N+1$ bidders participating in the auction. Building on the same reasoning as in 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, then they 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 all possible valuations that are below v^K.

For λ small enough, it cannot be optimal for j^* to fail to learn his value in such a way.

Hence, in any equilibrium that is robust to trembles, all buyers who learn $\max_j v_j = v^K$ and $v_i < v^K$ bid $\mathbb{E}[v_i|v_i < v^K] + u_i$.

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

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

The first line captures expected revenue when both the highest- and second-highest-valuation buyers have $v_i = v^K$. In such case the must 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

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

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

where

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

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

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

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

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

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

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.

$\square$

Proof of Proposition 4. Consider the following way to randomize access to the auction:

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

for some $q \in (0, 1)$. If only one buyer enters the auction $|M| = 1$, suppose he faces a random reserve price $\tilde{r}$ with support V. (When there is only one buyer in the auction, a reserve price is equivalent to a posted price.)

Take any information partition about self Π_λ^{self} in the support of the 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: $\Pi_\lambda^{self} = \{\{v\}_{v \in V}\}$. By contradiction, suppose not: Π^{self} bundles some values 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 $\pi_{\underline{v},\overline{v}}^{self}$ and let $\sigma_\lambda(\pi^{other}, \pi_{\underline{v},\overline{v}}^{self}, u_i)$ a buyer's equilibrium bid at that bundle.

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

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

However the finer partition must (weakly) increase buyer i's gross payoff. We show that it actually strictly increases i's gross payoff. Indeed, with probability $\Pr(\tilde{r} = \overline{v})q/N$, buyer i faces a posted price of $\overline{v}$, at which he wants to buy if $v_i = \overline{v}$ and $u_i > 0$, but not otherwise. Similarly, with probability $\Pr(\tilde{r} = \underline{v})q/N$, buyer i faces a posted price of $\underline{v}$, at which he does not want to buy if $v_i = \underline{v}$ and $u_i < 0$, but wants to otherwise. Hence the finer information partition increases i's gross payoff by at least

$$\frac{q}{N} \Pr(\tilde{r} = \overline{v}) \Bigg\{ \Pr\left(u_i > \overline{v} - \mathbb{E}\left(v_i \mid v_i \in \pi_{\underline{v},\overline{v}}^{self}\right)\right) p(\underline{v}) \left[\overline{v} - \underline{v} - \overline{u}\right]$$

$$+ \Pr\left(0 < u_i < \overline{v} - \mathbb{E}\left(v_i \mid v_i \in \pi_{\underline{v},\overline{v}}^{self}\right)\right) p(\overline{v}) \mathbb{E}\left[u_i \mid 0 < u_i < \overline{v} - \mathbb{E}\left(v_i \mid v_i \in \pi_{\underline{v},\overline{v}}^{self}\right)\right] \Bigg\}$$

$$+\frac{q}{N}\Pr(\tilde{r}=\underline{v})\Bigg\{\Pr\left(u_i<\underline{v}-\mathbb{E}\left(v_i\mid v_i\in\pi^{self}_{\underline{v},\overline{v}}\right)\right)p(\overline{v})\left[\overline{v}+\underline{u}-\underline{v}\right]$$

$$-\Pr\left(0>u_i>\underline{v}-\mathbb{E}\left(v_i\mid v_i\in\pi^{self}_{\underline{v},\overline{v}}\right)\right)p(\underline{v})\mathbb{E}\left[u_i\mid 0>u_i>\underline{v}-\mathbb{E}\left(v_i\mid v_i\in\pi^{self}_{\underline{v},\overline{v}}\right)\right]\Bigg\},$$

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

In any tremble-robust equilibrium, fully informed buyers must bid their valuations for the good: $\sigma((\pi^{other},\{v_i\},u_i))=v_i+u_i=\nu_i$. 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 Theorem 4. Consider the randomized access from Proposition 4. 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, the optimal reserve price r^* must target buyers with the highest possible realization of v_i, that is, $r^*=v^K+u^*$ for some $u^*\in(\underline{u},\overline{u})$. Suppose not, and $r^*<v^K+\underline{u}$ for all N, even as λ goes to zero. Using the same argument as in the proof of Theorem 3, we show that buyers must learn whether there toughest competitor has the highest valuation possible $\{v^K\}\in\Pi^{other}$ for N large enough. Suppose not, such that an equilibrium puts non-vanishing weight on $\Pi^{other}=\{\{v^{\underline{k}_l},\dots,v^{\overline{k}_l}\}_{l=1,\dots,L-1},\{v^{\underline{k}_L},\dots,v^K\}\}$ with $v^{\underline{k}_L}\leq v^{K-1}$. We know from Theorem 1 that after learning $\max_j v_j\in\{v^{\underline{k}_L},\dots,v^K\}$, they must choose $\Pi^{self}=\{\{v_i:v_i<v^{\underline{k}_L}\},\{v_i\}_{v_i\geq v^{\underline{k}_L}}\}$. Deviating to $\Pi^{other}=\{\{v^{\underline{k}_l},\dots,v^{\overline{k}_l}\}_{l=1,\dots,L-1},\{v^{\underline{k}_L},\dots,v^{K-1}\},\{v^K\}\}$ only increases their information costs by a vanishing amount for N large enough, while it induces a first-order reduction in their information costs on self. (See proof of Theorem 3 for more details.)

If a buyer learns that $\max_j v_j=v^K$, he then chooses $\Pi^{self}=\{\{v_i:v_i<v^K\},\{v^K\}\}$ and bids $\mathbb{E}[v_i\mid v_i<v^K]+u_i$ whenever $v_i<v^K$. A reserve price of $r<v^K+\underline{u}$ then

yields revenue strictly bounded above by

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

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

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

The latter is strictly greater whenever

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

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

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

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

This rewrites as

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

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. 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 $\mathbb{E}[\nu_{2:N} \mid \nu_{2:N} \leq v^K + u^*, \nu_{1:N} > v^K + u^*] - \varepsilon = \mathbb{E}[\nu_{1:N-1} \mid \nu_{1:N-1} \leq v^K + u^*] - \varepsilon$. 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

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

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

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

$$
- N\left(p^K \Pr(u_i \geq u^*)\right)\left[1 - \left(\frac{1-p^K}{1-p^K \Pr(u_i \geq u^*)}\right)^{N-1}\right]
$$

$$
\times \left(u^* - \mathbb{E}[u_{1:M} \mid M = |\{i : v_i = v^K, u_i < u^*\}|]\right)
$$

$$
- N\left(p^K \Pr(u_i \geq u^*)\right)\left(\frac{1-p^K}{1-p^K \Pr(u_i \geq u^*)}\right)^{N-1}
$$

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

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 $u^* + \underline{u}$ (or alternatively as long as the difference between u^* and the expected highest noise terms over all agents with $v_i = v^K$ and $u_i < u^*$ goes to zero sufficiently fast). □

B.5 Proof of Results of Section 6

Proof of Proposition 5. As in the proof of Proposition 1, we show that for λ small enough, there exists a tremble-robust symmetric equilibrium in which buyers become fully informed about their valuations $\Pi_N^{self} = \{\{v_i\}_{v_i \in V}\}$. Recall that buyers who fully learn their valuations must bid truthfully in any tremble-robust equilibrium, hence $\sigma_N(\{v_i\}, u_i) = v_i + u_i$. We can use the very same argument as in the proof of Proposition 1 to show that this constitutes an equilibrium for λ small enough. The only change in the proof is that an buyer i's belief about others' values (and hence bids) now depends on what he learns about himself, as valuations are correlated across buyers through their dependence on ω. However, since p_ω has full support for all ω, we simply need to replace prior probabilities p with i's posterior about $\max_j v_j$ given his information π_i, and the proof goes through.

We now need to show that such equilibrium fully reveals the common component at $N \longrightarrow \infty$. Recall that buyers always learn their true valuations and bid truthfully: $(b_i)_i = (v_i + u_i)_i$. Given that the noise terms are small, by observing $b_i = v_i + u_i$ the seller can always infer v_i perfectly, which is the only part of the bid that is informative of ω. The seller's posterior given $(v_i)_i$ is then

$$
\Pr(\omega \mid (v_i)_i) = \frac{\mu_0(\omega) \prod_i p_\omega(v_i)}{\sum_{\omega'} \mu_0(\omega') \prod_i p_{\omega'}(v_i)}.
$$

We want to show that, for any $\varepsilon > 0$,

$$\lim_{N \longrightarrow \infty} \Pr(||\tilde{\mu}_N - \delta_{\tilde{\omega}}|| \geq \varepsilon) = 0.$$

We show that

$$\lim_{N \longrightarrow \infty} \Pr(||\tilde{\mu}_N - \delta_{\omega}|| \geq \varepsilon \mid \omega) = 0 \quad \forall \omega \in \Omega,$$

which implies the above by the law of total probability. Recall that δ_ω is the belief that puts probability one on the common component being ω, i.e. $\delta_\omega(\omega') = \mathbb{1}\{\omega' = \omega\}$. Note that if $\mu_N(\omega) > 1 - \varepsilon/\sqrt{2}$, then

$$\begin{aligned}||\mu_N - \delta_\omega|| = \sqrt{(1-\mu_N(\omega))^2 + \sum_{\omega' \neq \omega} \mu_N(\omega')^2} &\leq \sqrt{(1-\mu_N(\omega))^2 + (1-\mu_N(\omega))^2} \\ &< \sqrt{2\left(\frac{\varepsilon}{\sqrt{2}}\right)^2} = \varepsilon.\end{aligned}$$

Hence it suffices to show that $\Pr\left(\mu_N(\omega) > 1 - \varepsilon/\sqrt{2} \mid \omega\right) \longrightarrow 1$ as $N \longrightarrow \infty$.

Let $\hat{\varepsilon} \equiv \varepsilon/\sqrt{2}$. We want to show that $\mu_N(\omega) > 1 - \hat{\varepsilon}$ with a probability that tends to one. A sufficient condition is then that

$$\frac{\mu_N(\omega')}{\mu_N(\omega)} < \frac{1}{|\Omega| - 1} \frac{\hat{\varepsilon}}{1 - \hat{\varepsilon}} \quad \forall \omega' \neq \omega.$$

Indeed, summing over all $\omega' \neq \omega$ this implies

$$\sum_{\omega' \neq \omega} \frac{\mu_N(\omega')}{\mu_N(\omega)} < \frac{\hat{\varepsilon}}{1 - \hat{\varepsilon}} \iff \frac{1 - \mu_N(\omega)}{\mu_N(\omega)} < \frac{\hat{\varepsilon}}{1 - \hat{\varepsilon}} \iff \mu_N(\omega) > 1 - \hat{\varepsilon}.$$

So we have to show that $\Pr\left(\frac{\mu_N(\omega')}{\mu_N(\omega)} < \frac{1}{|\Omega|-1} \frac{\hat{\varepsilon}}{1-\hat{\varepsilon}} \forall \omega' \neq \omega \mid \omega\right) \longrightarrow 1$ as $N \longrightarrow \infty$.

Since buyers learn their valuations fully and bid truthfully, the distribution of bids conditional on a realization of ω converges in probability to the distribution of valuations p_ω, by the weak law of large numbers. That is, for every $\eta > 0$, there exists $\overline{N}$ such that for all $N \geq \overline{N}$,

$$\Pr\left((v_i)_{i=1}^N \text{ s.t. } \frac{|\{i \mid v_i = v\}|}{N} \in [p_\omega(v) - \eta, p_\omega(v) + \eta] \; \forall v \in V\right) \geq 1 - \eta.$$

In words, the probability of observing a vector of values such that a proportion of buyers with value v is very close to $p_\omega(v)$ goes to one. When the seller observes such vector of bids (or equivalently, vector of values), his posterior satisfies

$$\frac{\mu_N(\omega')}{\mu_N(\omega)} = \frac{\mu_0(\omega')}{\mu_0(\omega)} \prod_{i=1}^{N} \frac{p_{\omega'}(v_i)}{p_\omega(v_i)} = \frac{\mu_0(\omega')}{\mu_0(\omega)} \prod_{v\in V} \left(\frac{p_{\omega'}(v)}{p_\omega(v)}\right)^{|\{i|v_i=v\}|},$$

which, by construction, lies in

$$\begin{aligned}\frac{\mu_N(\omega')}{\mu_N(\omega)} &\in \left[\frac{\mu_0(\omega')}{\mu_0(\omega)} \prod_{v\in V} \left(\frac{p_{\omega'}(v)}{p_\omega(v)}\right)^{N(p_\omega(v)-\eta)}, \frac{\mu_0(\omega')}{\mu_0(\omega)} \prod_{v\in V} \left(\frac{p_{\omega'}(v)}{p_\omega(v)}\right)^{N(p_\omega(v)+\eta)}\right] \\ &= \left[\frac{\mu_0(\omega')}{\mu_0(\omega)} \left(\prod_{v\in V} \left(\frac{p_{\omega'}(v)}{p_\omega(v)}\right)^{p_\omega(v)-\eta}\right)^N, \frac{\mu_0(\omega')}{\mu_0(\omega)} \left(\prod_{v\in V} \left(\frac{p_{\omega'}(v)}{p_\omega(v)}\right)^{p_\omega(v)+\eta}\right)^N\right].\end{aligned}$$

Furthermore

$$\log\left(\prod_{v\in V} \left(\frac{p_{\omega'}(v)}{p_\omega(v)}\right)^{p_\omega(v)}\right) = \sum_v p_\omega(v) \log\left(\frac{p_{\omega'}(v)}{p_\omega(v)}\right) < \log\left(\sum_v p_\omega(v) \frac{p_{\omega'}(v)}{p_\omega(v)}\right) = 0,$$

by Jensen's inequality since $\log$ is a strictly concave function. Hence, for η small enough,

$$\prod_{v\in V} \left(\frac{p_{\omega'}(v)}{p_\omega(v)}\right)^{p_\omega(v)-\eta} < 1 \quad \text{and} \quad \prod_{v\in V} \left(\frac{p_{\omega'}(v)}{p_\omega(v)}\right)^{p_\omega(v)+\eta} < 1$$

such that $\frac{\mu_N(\omega')}{\mu_N(\omega)} \longrightarrow 0$ as $N \longrightarrow \infty$ whenever the empirical distribution of values is within η of the true distribution p_ω. Since η can be arbitrarily close to zero, such event can have a probability arbitrarily close to one: $\Pr\left(\frac{\mu_N(\omega')}{\mu_N(\omega)} < \frac{1}{|\Omega|-1}\frac{\hat{\varepsilon}}{1-\hat{\varepsilon}} \forall \omega' \neq \omega \mid \omega\right) \longrightarrow 1$ as $N \longrightarrow \infty$. □

Proof of Proposition 6. **Step 0.** We first argue that our analysis of symmetric, tremble-robust equilibria in the proof of Theorem 2 still holds in this extended environment where valuations are correlated across buyers through their dependence on ω. A key reason is that $\operatorname{supp} p_\omega = V$ for all ω, and so learning something about the competitor's value $\max_j v_j$ can never lead an agent to rule out *for sure* some values he might have. Indeed, if agent i learns that $\max_j v_j \in \pi^{other}$, that does provide him with some information on v_i but his posterior about v_i still has full support:

$$\begin{aligned}\Pr\left(v_i = v \mid \max_j v_j \in \pi^{other}\right) &= \sum_\omega \Pr\left(\omega \mid \max_j v_j \in \pi^{other}\right) p_\omega(v) \\ &= \sum_\omega \frac{\Pr\left(\max_j v_j \in \pi^{other} \mid \omega\right) \mu_0(\omega)}{\Pr\left(\max_j v_j \in \pi^{other}\right)} p_\omega(v) > 0\end{aligned}$$

since $p_\omega(v) > 0$ for all v, ω, and similarly $\Pr\left(\max_j v_j \in \pi^{other} \mid \omega\right) > 0$. Hence even if learning about others' valuations provides i with some information about his own, there must always remain some full support uncertainty. For λ small enough, agent i will then want to resolve that uncertainty whenever there are strict gains from doing so. Agent i will still not find it worthwhile to learn to distinguish all the values he might have that lie for sure below his toughest competitor's.

Step 1. We now show that for N large enough, buyers must find it optimal to learn whether their toughest competitor has the highest possible valuation: there exists $\overline{N}$ such that for all $N \geq \overline{N}$, there exists $\bar{\lambda}$ such that, for all $\lambda \leq \bar{\lambda}$, $\{v^K\} \in \Pi_N^{other}$. The proof is almost identical to the proof of Theorem 3 (Step 1), except that we now have to account for the fact that when i learns something about his competitors $\max_j v_j$, it also gives him information about his own valuation.

By contradiction, suppose not: for all N, Π_N^{other} bundles v^K with some lower values. Consider an alternative strategy $\widehat{\Pi}_N^{other}$ that is the same as Π_N^{other} except for the fact that it unbundles v^K. This alternative partition is strictly costlier, since it is strictly finer than the one we started with. However $c\left(\widehat{\Pi}_N^{other}, p_{1:N-1}\right) - c\left(\Pi_N^{other}, p_{1:N-1}\right)$ becomes arbitrarily small as N grows large (see the proof of Theorem 3 for more details).

There is however a first-order gain in choosing this finer partition $\widehat{\Pi}_N^{other}$ as it allows buyers to save on information cost about their own valuations. If buyers learn that $\max_j v_j \in \{v^{\underline{k}_L}, \ldots, v^K\}$, then they optimally chooses to partition their own set of valuations into $\{\{v^1, \ldots, v^{\overline{k}_L - 1}\}, \{v^k\}_{k=\underline{k}_L,\ldots,K}\}$. The same is true if buyers learn that $\max_j v_j \in \{v^{\underline{k}_L}, \ldots, v^{K-1}\}$. If however buyers learn that $\max_j v_j = v^K$, then they optimally choose a coarser partition about their own valuations: $\{\{v^1, \ldots, v^{K-1}\}, \{v^K\}\}$. So far, the argument is the same as in the proof of Theorem 3. The difference stems from the fact that learning about $\max_j v_j$ changes i's belief about v_i, and so the cost of a partition on self is no longer $c(\Pi^{self}, p)$ but $c(\Pi^{self}, \Pr_{v_i}(\cdot \mid \pi_i^{other}))$. The reduction in information cost on self associated with $\widehat{\Pi}_N^{other}$ is then

$$
\begin{aligned}
&\Pr\left(\max_j v_j = v^K\right)\Bigg[c\left(\{\{v^1,\dots,v^{K-1}\},\{v^K\}\},\Pr(\cdot\mid \max_j v_j = v^K)\right) \\
&\qquad\qquad - c\left(\{\{v^1,\dots,v^{\tau_L}\},\{v^k\}_{k=\tau_L+1}^K\},\Pr(\cdot\mid v^{\underline{k}_L}\le \max_j v_j \le v^K))\right)\Bigg] \\
&+\Pr\left(v^{\underline{k}_L}\le \max_j v_j < v^K\right)\Bigg[c\left(\{\{v^1,\dots,v^{\overline{k}_L-1}\},\{v^k\}_{k=\underline{k}_L}^K\},\Pr(\cdot\mid v^{\underline{k}_L}\le \max_j v_j \le v^{K-1})\right) \\
&\qquad\qquad - c\left(\{\{v^1,\dots,v^{\overline{k}_L-1}\},\{v^k\}_{k=\underline{k}_L}^K\},\Pr(\cdot\mid v^{\underline{k}_L}\le \max_j v_j \le v^K)\right)\Bigg].
\end{aligned}
$$

Since $\Pr(\max_j v_j = v^K)$ tends to one as N tends to infinity, the second term goes away for N large enough and we can focus on the first one. We need to show that the first term remains strictly positive (and does not go to zero) as N goes to infinity. To distinguish between the change in partition on self, and the change in belief about self, we can rewrite that term as $\Pr\left(\max_j v_j = v^K\right)$ times

$$
\begin{aligned}
&c\left(\{\{v^1,\dots,v^{K-1}\},\{v^K\}\},\Pr(\cdot\mid \max_j v_j = v^K)\right) \\
&\qquad\qquad - c\left(\{\{v^1,\dots,v^{K-1}\},\{v^K\}\},\Pr(\cdot\mid v^{\underline{k}_L}\le \max_j v_j \le v^K))\right) \\
&+ c\left(\{\{v^1,\dots,v^{K-1}\},\{v^K\}\},\Pr(\cdot\mid v^{\underline{k}_L}\le \max_j v_j \le v^K))\right) \\
&\qquad\qquad - c\left(\{\{v^1,\dots,v^{\tau_L}\},\{v^k\}_{k=\tau_L+1,\dots,K}\},\Pr(\cdot\mid v^{\underline{k}_L}\le \max_j v_j \le v^K))\right).
\end{aligned}
$$

The difference between the last two terms must be strictly negative, as we are comparing the cost of two partitions, one strictly finer than the other, at a given prior belief. The difference between the first two terms must go to zero as N goes to infinity, as we are comparing the cost of the same partition, but just changing buyer i's prior about v_i from $\Pr(\cdot \mid \max_j v_j = v^K)$ to $\Pr(\cdot \mid v^{\underline{k}_L} \le \max_j v_j \le v^K))$. However, as N grows large, buyer i puts a weight that tends to one on the event $\max_j v_j = v^K$ when he learns $v^{\underline{k}_L} \le \max_j v_j \le v^K$. Hence these two beliefs effectively carry the same information about ω, and hence about v_i, for N sufficiently large.

It must be that for N large enough, the gains associated with cheaper information acquisition on self more than compensate the cost of the extra information about others, and $\widehat{\Pi}_N^{other}$ yields a strictly higher payoff to the buyer than Π_N^{other}.

Step 2. We construct an environment $\{p_\omega\}_\omega$ in which the auction fails to reveal the common component as N tends to infinity. In particular, we show that this is the case whenever $p_\omega(v^K) = p_{\omega'}(v^K)$ for some $\omega \neq \omega'$.

We know from Step 1 that, if N is sufficiently large, buyers learn whether their toughest competitor has a value $\max_j v_j = v^K$, and if yes, choose to bundle all the values they can have that are below: $\Pi^{self} = \{\{v^1, \ldots, v^{K-1}\}, \{v^K\}\}$. Of course, as N goes to infinity, the probability that at least two buyers have $v_i = v^K$ goes to one. In such event, all buyers with $v_i = v^K$ learn their valuations fully and bid it: $\sigma_N(\pi = (\{v^K\}, \{v^K\}, u_i)) = v_i + u_i$. All others however fail to learn their valuations and bid

$$\sigma_N(\{v^K\}, \{v^1, \ldots, v^{K-1}\}, u_i) = \mathbb{E}\left[v_i \mid v_i \leq v^{K-1}, \max_j v_j = v^K\right] + u_i.$$

The only information about buyers' valuations that is captured by the bids is then the share of buyers with $v_i = v^K$. This is not enough to learn about the common component if $p_\omega(v^K) = p_{\omega'}(v^K)$ for some $\omega \neq \omega'$. Indeed, the seller's posterior satisfies

$$\begin{aligned}
\frac{\mu_N(\omega')}{\mu_N(\omega)} &= \frac{\mu_0(\omega')}{\mu_0(\omega)} \prod_{i=1}^{N} \frac{\Pr(\sigma_N(\pi_i)|\omega')}{\Pr(\sigma_N(\pi_i)|\omega)} \\
&= \frac{\mu_0(\omega')}{\mu_0(\omega)} \left(\frac{p_{\omega'}(v^K)}{p_\omega(v^K)}\right)^{|\{i|\sigma_N(\pi_i)=v^K+u_i\}|} \left(\frac{1-p_{\omega'}(v^K)}{1-p_\omega(v^K)}\right)^{|\{i|\sigma_N(\pi_i)<v^K+\underline{u}\}|} \\
&= \frac{\mu_0(\omega')}{\mu_0(\omega)},
\end{aligned}$$

for all $N \geq \overline{N}$. □

B.6 Proof of Results in the Additional Material

Proof of Proposition 7. 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, the signal about self Π^{self} that buyers acquire must fully reveal v_i. Suppose not, such that Π^{self} bundles two possible values (i.e., there exists $\pi^{self} \in \Pi^{self}$ such that $|\pi^{self}| > 1$) and let v_i' and v_i'' 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 can be at the same information set about self 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 $v_i' + u_i$ and $v_i'' + u_i$, and given that they face such bid with non-vanishing probability, they have a strict incentive to learn to distinguish value v_i' from value v_i''.

If a buyer learns his value fully $\Pi^{self} = \{\{v_i\}_{v_i \in V}\}$, then information about others has strictly no value, and so he must choose $\Pi^{other} = \{V\}$. We however know from Proposition 2 that such an equilibrium, in which buyers converge to becoming fully informed of their valuations, cannot exist. $\square$