Information Structures in Optimal Auctions

by

Dirk Bergemann and Martin Pesendorfer

August 2001

COWLES FOUNDATION DISCUSSION PAPER NO. 1323

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY

Box 208281

New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Information Structures in Optimal Auctions*

Dirk Bergemann[†] Martin Pesendorfer[‡]

First Version: June 2000
This Version: August 2001

Abstract

A seller wishes to sell an object to one of multiple bidders. The valuations of the bidders are privately known. We consider the joint design problem in which the seller can decide the accuracy by which bidders learn their valuation *and* to whom to sell at what price. We establish that optimal information structures in an optimal auction exhibit a number of properties: (i) information structures can be represented by monotone partitions, (ii) the cardinality of each partition is finite, (iii) the partitions are asymmetric across agents. These properties imply that the optimal selling strategy of a seller can be implemented by a sequence of exclusive take-it or leave-it offers.

KEYWORDS: Optimal Auction, Private Values, Information Structures, Partitions.
JEL CLASSIFICATION: C72, D44, D82, D83.

*The authors thank Nicolas Hengartner, Bill Sudderth, Marten Wegkamp and Steve Williams for several helpful discussions. We thank participants at a seminar at the University of Illinois for useful comments. Financial support from NSF Grant SES 0095321 and a Sloan Research Fellwoship, and NSF Grant SES 9811134, respectively, is gratefully acknowledged.

[†]Department of Economics, Yale University, New Haven, CT 06511, dirk.bergemann@yale.edu.
[‡]Department of Economics, Yale University, New Haven, CT 06511, martin.pesendorfer@yale.edu.

1 Introduction

The optimal design of an auction has received considerable attention in the economics literature. Applications of the literature arise in variety of settings including the sale of publicly and privately owned companies and the sale of radio spectrum licenses. Myerson (1981) constitutes the seminal paper in the field. Myerson shows which auction rules achieve the largest revenues to the seller in a single object auction. Most of the subsequent literature on mechanism design maintains the assumption that the information held by market participants is given as exogenous. Little is known about optimal mechanisms when the information of the participants is allowed to be endogenous.

This paper considers the optimal auction design problem when the seller can decide how much information is revealed to individual bidders. Examples in which the amount of information available to the buyers is determined by the seller are numerous. The sale of a company in which the selling party may release proprietary financial or product information to one or more bidders is one prominent example. In these auctions the nature of the information released to the bidders may differ across bidders. The sale of a real estate property in which the seller shows the property to the buyers individually is another example. The seller can then decide which attributes of the house to emphasize during a showing, and which parts to de-emphasize. What is common in these examples is that the seller chooses the accuracy of the buyers information.

We consider a problem in which a seller offers a single object to a number of risk neutral bidders. The seller wishes to maximize revenues from the sale. Bidders' valuations for the object are private and not known prior to the bidding. The seller controls the bidders' information structures which generate the bidders' private information. The information structure determines the accuracy with which buyers learn their valuations prior to the auction. The seller may assign an information structure that informs a bidder perfectly or an information structure that gives the bidder only a rough guess about her true value for the object.

The role of the information structure is easily illustrated with the sale of a company. Suppose the primary value of the company, the target company, to be sold is its client list. The value of the target to each bidding firm is determined by the match of its own client list with the one of the target company. The bidding company may value new clients which would be added to its existing list of clients, or it may value clients appearing on both lists for the purpose of cross-selling. The bidding firm can improve its estimate about the value of the target company if it receives a detailed description of the client list of the target company. Thus, the selling party can influence the quality of the estimate of the bidders by releasing more or less detailed information about the client list. However, the selling party, despite the perfect control it has over the release of information, will never know the private value the bidding company eventually attaches to the target company, as the seller doesn't observe the bidding company's client list. In other words, the seller selects the information structure for the bidder, but does not observe the private value realization of the bidder.

In general terms our model encompasses all situations where the preferences of the bidders are privately known to the bidders and the characteristics of the object are proprietary information of the seller. The proprietary information can be released more or less complete by the seller. Provided the characteristics of the object matter for the bidders' valuations the choice of information structure will influence how the preferences are mapped into valuations for the object.

After the choice of information structure by the seller, the bidders then report their value estimate to a revelation mechanism which determines the probability of winning the object and a transfer payment for every bidder. We study information structures and revelation mechanisms that maximize the seller's revenues. The solution in Myerson (1981) arises in our model as a special case when the seller informs the bidders perfectly.

Milgrom & Weber (1982) consider a related problem in which the seller can decide whether to release information publicly. They consider an informational environment in which bidders valuations are affiliated with a single object in first price, second price and English auctions. They show that the seller's revenues can only increase by releasing information to all bidders, which is the so-called linkage principle. The intuition is that losing bids underestimate other bidders' signals. Losing bids rise on average (and therefore so too do revenues) when the seller reveals his information. Perry & Reny (1999) show that the linkage principle may fail in other auctions.

In contrast to the problem analyzed by Milgrom and Weber we consider an independent private valuations framework. We assume that the seller can choose the accuracy of the buyers information without knowing the information of the buyers. Thus, the seller's choice can affect buyers differentially. In contrast, in the Milgrom and Weber model information is released publicly *and* affects buyers symmetrically.

The endogenous acquisition of information in principal agent problems has been studied by a number of authors: Cremer & Khalil (1992), Sobel (1993), Lewis & Sappington (1993), Cremer & Rochet (1998*a*), Cremer & Rochet (1998*b*) study incentive contracts in which the precision of private information by the agent is endogenous. In these single agent models, the agent rather than the principal controls the acquisition of information. The role of endogenous information in many agents models has been investigated in the private values setting by Tan (1992) and Stegeman (1996). Again, the focus in these papers is on the decisions of the agents rather than the principal's control.[1]

[1] As for interdependent and common value auctions, Matthews, (1977) and (1984), considers endogenous informa-

The endogenous choice of the precision of information in revenue maximizing auctions is motivated by two opposing effects of information: First, more information increases the efficiency of the auction and thus seller's revenues. Second, more information increases the rents of the bidders in form of information rents which lower the seller's revenues. We analyze this trade-off and characterize the properties of optimal information structures under mild regularity assumptions on initial informational priors.

Our paper is organized as follows: Section 2 describes the model. Section 3 considers two examples. The first example illustrates the (special and perhaps trivial) case of an auction with one bidder. The second example considers an auction with two bidders in which valuations are drawn from the uniform distribution function. We illustrate that the optimal information structure involves at most a binary partition and is asymmetric. It has the feature that one bidder learns whether her valuation is high or low and the other bidder does not learn anything about her valuation. We then briefly argue how the insights of these examples generalize to auctions with many bidders.

Section 4 establishes that the revenues of the seller are maximized with discrete information structures. Locally, for any bidder and for any point of continuity in the distribution function of the bidder, there is a benefit of introducing a mass point into the distribution function. There are two opposing effects: First, truthful type reports are more costly with continuous than with discrete probability distributions of types. Second, discrete signals about valuations are inaccurate and there will be a loss in revenues due to the induced efficiency loss. We show that locally the first effect dominates.

Section 5 first summarizes results of the optimal auction design problem when types are drawn from a discrete probability distribution function. Based on these results we then study the optimal information structure in an optimal auction. We show the following results: (i) optimal information structure are partitions where each partition can be represented by a countable collection of disjoint intervals that cover the space of valuations, (ii) the number of partitions is finite, (iii) optimal partitions exist.

Section 6 establishes additional results. We show that the number of elements in the partition increases as the number of bidder increases. Finally, we point out that the optimal auction under the optimal information structure can be implemented by a sequence of take-it or leave-it offers. The auctioneer starts at a high price and asks the first bidder whether she is willing to buy the object. If the offer is rejected the auctioneer moves on to the next bidder and lowers the asking

tion acquisition in a pure *common* values auctions. Persico (2000) compares the equilibrium incentives of the bidders to acquire information in first and second price auctions within a model of affiliated values. Finally, Bergemann & Valimaki (forthcoming) consider information acquisition in a general efficient mechanism design setting.

price. After all the bidders have been asked, the auctioneer returns to the first bidder, but with a lower asking price and the process repeats. The maximal number of take it or leave it offers is finite. The implementation result follows from two properties of optimal partitions: First, optimal virtual utilities are strictly increasing for any initial informational distribution. Second, optimal partitions induce finite and asymmetric partitions for bidders although initial prior distributions are permitted to be symmetric. Section 7 concludes.

2 Model

Subsection 2.1 specifies the auction environment. Subsection 2.2 introduces in some detail the notion of an information structure and related concepts. Finally, Subsection 2.3 defines the notion of a direct revelation mechanism and associated constraints.

2.1 Utility

A seller has a single object for sale. There are I potential bidders for the auction, indexed by $i \in \{1, ..., I\}$. Each agent i has a compact set $V_i = [0, 1]$ of possible valuations for the object, where a generic element is denoted by $v_i \in V_i$, and

$$V = \mathop{\times}_{i=1}^{I} V_i = [0, 1]^I .$$

We occasionally adopt the notation $v = (v_i, v_{-i})$. The valuation v_i is independently distributed with prior distribution function $F_i(v_i)$. The prior distribution function $F_i(v_i)$ is common knowledge. The associated density function $f_i(v_i)$ is positive and differentiable on V_i. The utility of the agent is quasilinear and given by

$$u_i(v_i, t_i) = v_i - t_i,$$

where t_i is a monetary transfer.

2.2 Information Structure

The signal space S_i is compact and without loss of generality $S_i \subseteq V_i$. The space S_i can either be countable, finite or infinite, or uncountable. Let $(V_i \times S_i, \mathcal{B}(V_i \times S_i))$ be a measurable space, where $\mathcal{B}(V_i \times S_i)$ is the class of Borel sets of $V_i \times S_i$. An *information structure* for agent i is given by a pair $\langle S_i, F_i(v_i, s_i)\rangle$, where S_i is the space of signal realizations and $F(v_i, s_i)$ is a joint probability distribution over the space of valuations V_i and the space of signals S_i. The joint probability distribution is defined in the usual way by

$$F_i(v_i, s_i) = \Pr(v \le v_i, s \le s_i).$$

The marginal distributions of $F_i(v_i, s_i)$ are denoted with minor abuse of notation by $F_i(v_i)$ and $F_i(s_i)$ respectively. For $F_i(v_i, s_i)$ to be part of an information structure requires that the marginal distribution with respect to v_i equals the prior distribution over v_i. The conditional distribution functions derived from joint distribution function are defined in the usual way by:

$$F_i(v_i | s_i) \triangleq \frac{\int_0^{v_i} dF_i(\cdot, s_i)}{\int_0^1 dF_i(\cdot, s_i)}$$

and similarly,

$$F_i(s_i | v_i) \triangleq \frac{\int_0^{s_i} dF_i(v_i, \cdot)}{\int_0^1 dF_i(v_i, \cdot)}.$$

The auctioneer can choose an arbitrary information structure $\langle S_i, F_i(v_i, s_i)\rangle$ for every bidder i subject only to the restriction that the marginal distribution equals the prior distribution of v_i. The cost of every information structure is identical and set equal to zero. The choice of $\langle S_i, F_i(v_i, s_i)\rangle$ is common knowledge among the bidders. At the interim stage every agent observes privately a signal s_i rather than her true valuation v_i of the object. Given the signal s_i and the information structure $\langle S_i, F_i(v_i, s_i)\rangle$ each bidder forms an estimate about her true valuation of the object. The expected value of v_i conditional on observing s_i is given by

$$\mathbb{E}[v_i | s_i] = \int_0^1 v_i dF_i(v_i | s_i),$$

and to distinguish the posterior realizations of expected values from the prior values, we denote

$$w_i(s_i) \triangleq \mathbb{E}[v_i | s_i].$$

Every information structure $\langle S_i, F_i(v_i, s_i)\rangle$ generates a distribution function $G_i(w_i)$ over posterior expectations, which is given by

$$G_i(w_i) = \int_{\{s_i : w_i(s_i) \le w_i\}} dF_i(s_i).$$

We denote by W_i the support of the distribution function $G_i(\cdot)$. We impose the following regularity conditions: (i) $G_i(w_i)$ can be represented by a convex combination of an absolute continuous function (with respect to the Lebesgue measure) and a discrete function; (ii) the density $g_i(w_i)$ in the absolute continuous part is once differentiable.[2]

[2] The regularity conditions are used only in the proof of Theorem 2. Theorem 7, which shows the optimality of a finite information structure, does not rely on these regularity assumptions. The proof of Theorem 7 could be extended to give a separate proof of Theorem 2 without the regularity conditions, in particular, the continuity and differentiability assumptions.

Observe that the prior distribution $F_i(\cdot)$ and the posterior distribution over expected values $G_i(\cdot)$ need not to coincide. For future discussions it is helpful to illustrate some specific information structures. The information structure $\langle S_i, F_i(v_i, s_i)\rangle$ yields perfect information if $F_i(v_i) = G_i(v_i)$ for all $v_i \in V_i$. In this case, the conditional distribution $F(s_i | v_i)$ has to satisfy

$$F(s_i | v_i) = \begin{cases} 0 & \text{if} \quad s_i < s(v_i), \\ 1 & \text{if} \quad s_i \geq s(v_i), \end{cases} \tag{1}$$

where $s(v_i)$ is an invertible function. An information structure which satisfies (1) without necessarily satisfying the invertibility condition is called *partitional*.[3] An information structure is called *discrete* if S_i is countable and *finite* if S_i is finite.

After the choice of the information structures $\langle S_i, F_i(v_i, s_i)\rangle$ by the auctioneer, the induced distribution of the agents valuation is given by $G_i(w_i)$ rather than $F_i(v_i)$. The signal s_i and the corresponding expected valuation $w_i(s_i)$ remain private signals for every agent i and the auctioneer still has to elicit information by respecting the truthtelling conditions.

2.3 Mechanism

The seller selects the information structures of the bidders and a revelation mechanism. By the revelation principle we may restrict attention to direct revelation mechanism. *A direct revelation mechanism* consists of a tuple $(W_i, t_i, q_i)_{i=1}^I$ with the transfer payment of bidder i:

$$t_i : \bigtimes_{i=1}^{I} W_i \to \mathbb{R},$$

and the probability of winning the object for bidder i:

$$q_i : \bigtimes_{i=1}^{I} W_i \to [0, 1].$$

We sometimes write $T_i(w_i)$ for the expected transfer payment,

$$T_i(w_i) \triangleq \mathbb{E}_{w_{-i}} t_i(w_i, \cdot),$$

where the expectation is taken over $w_{-i} = (w_1, ..., w_{i-1}, w_{i+1}, ..., w_I)$. Similarly, $Q_i(w_i)$ denotes the expected probability of winning,

$$Q_i(w_i) \triangleq \mathbb{E}_{w_{-i}} q_i(w_i, \cdot).$$

[3] A common and equivalent representation of information structures, see Laffont (1989), Chapter 4, is to start with a prior distribution over V_i and then specify the conditional probability distribution over signals by a mapping ϕ_i, where $\phi_i : V_i \to \Delta(S_i)$. A *partitional* information structure or information structure *without noise* in this framework is given by a mapping $\phi_i : V_i \to S_i$. This and our approach are equivalent since the joint probability distribution defines a conditional distribution and conversely every conditional distribution together with a marginal distribution induces a joint distribution.

The interim utility of bidder i with an expected valuation w_i and announced valuation $\widehat{w}_i$ is:

$$U_i\left(w_i, \widehat{w}_i\right) = w_i Q_i(\widehat{w}_i) - T_i(\widehat{w}_i).$$

The mechanism has to satisfy the participation constraints:

$$U_i\left(w_i, w_i\right) \geq 0, \text{ for all } w_i \in W_i,$$

and the incentive compatibility constraints:

$$U_i\left(w_i, w_i\right) \geq U_i\left(w_i, \widehat{w}_i\right), \text{ for all } w_i, \widehat{w}_i \in W_i.$$

3 Examples

This section illustrates properties of optimal information structures for some special cases. First, we look at a single bidder auction. Second, we discuss properties of optimal information structures in an auction with two bidders. Third, we depict properties of the numerical solution to an auction with many bidders when the valuations are drawn from the uniform distribution. These examples illustrate that the seller prefers to reveal sparse information and treat bidders asymmetrically.

Consider first the case with a single bidder. The information structure in which the seller reveals all the information to the bidder is analyzed in Myerson (1981). Suppose for simplicity that the seller values the object at zero. Myerson establishes that the seller can extract at most the virtual valuations in any incentive compatible selling mechanism. The virtual valuation of a bidder of type v equals the type of the bidder minus the incentive cost,

$$v - \frac{1 - F(v)}{f(v)}$$

Notice, that the incentive cost is positive for and remains positive even if the seller reveals information only partially.

In contrast consider the situation in which the seller chooses to reveal no information at all to the buyer. Without any information a bidder is willing to pay up to the ex ante expected valuation of the bidder to receive the object. In this case, the seller can extract all the expected surplus. It is therefore immediate that revealing no information is optimal in a single bidder auction. The seller can post a price equal to the ex ante expected valuation. This posted price scheme extracts the total surplus *and* is efficient. Moreover, if the seller were to reveal some information to the bidder, the seller would be worse off because he incurs an incentive cost expressed by the virtual utility.

Next, suppose we add a second bidder with an identical prior distribution to the auction. The policy to disclose no information does not remain optimal with two bidders. To see this, notice that revealing no information extracts at most the ex ante expected valuation of the winning bidder. But with symmetric bidders, the revenue for the auctioneer would then be the same as in the case of a single bidder. In a two bidder auction there is a simple scheme that achieves more rent by exploiting the increase in the number of bidders. The scheme has the following feature: The seller discloses no information to the first bidder as in the case of a single bidder auction, but assigns a binary information structure to the second bidder. A binary information structure permits the bidder to determine whether the valuation is above or below a certain threshold. The optimal threshold equals the ex-ante expected value of the object. The scheme then works as follows: Initially, the seller offers the object to the second bidder at a price equal to the conditional expected valuation in the event that the valuation is above the threshold. If the second bidder rejects the offer, then the seller offers the object to the first bidder at a price equal to the ex ante expected valuation. The total revenues to the seller under this scheme exceed the ex ante expected valuation of a bidder. Thus, the revenues under this scheme are higher than under a scheme in which the seller reveals no information. We observe that as before the seller leaves no informational rent to the bidders. However, the allocation is not necessarily efficient anymore, as it could be that the first bidder has a higher valuation for the object than the second bidder. The coarseness of the information structure may prevent the seller to make the efficient choice.

In fact, it can be shown that the described information structure maximizes the revenues to the seller with two bidder and uniformly distributed valuations. If attention is restricted to the class of information structures with finite partitions, then this result follows immediately from the first and second order conditions for optimally chosen partitions. Our results in the subsequent sections establish that the described scheme with two bidders is indeed optimal for the uniform distribution under general information structures even permitting non-finite and non-partitional information structures.

The scheme with two bidders has a number of features that are worth emphasizing: First, even if bidders have initially symmetric prior distributions of valuations, they are optimally assigned asymmetric information structures. The first bidder receives no information, while the second bidder learns wether the valuation is above or below the ex-ante mean. Second, the seller does not give an informational rent to buyers. Both bidders are offered the object at a fixed price that they can accept or reject. Third, the information structure can be implemented by a sequence of take-it or leave-it offers.

A natural question is whether the features of the optimal information structure for two bidders with uniformly distributed valuations extend to more general settings. We address this question in the subsequent sections. Before we start our formal analysis we illustrate graphically optimal information structures with many bidders and uniform distributed valuations. The following figures depict properties of optimal information structures as we vary the number of bidders.

Insert Figure 1 about here

Figure 1 depicts the optimal partition structure for the bidder with the highest virtual valuation. As can be seen in the figure the number of partitional elements increases monotone with the number of bidders participating in the auction. However, the increase is only very gradual. We count three partitions with three to six bidders, four partitions with seven to fifteen bidders and five partitions with sixteen or more bidders.

Insert Figure 2 about here

Figure 2 depicts the positive virtual valuations for the bidder with partitions depicted in Figure 1. The figure illustrates that the highest virtual valuation increases, as the number of bidders increases. The number of virtual valuations depicted in the figure corresponds to the number of partitions in Figure 1 minus one. The figures illustrate further that in general it is not the case that the seller leaves no informational rent to the bidder. With three or more bidders binary partitions are no longer optimal and as the auctioneer has to reward agents to report truthfully, he will have to incur incentive costs. As the number of bidders increases, the information structure becomes finer. The intuition is that with more competition the incentive costs due the informational rents are lower and the revenue gains from improving allocative efficiency due to more information become more important, as the number of bidders increases.

Insert Figure 3 about here

Figure 3 depicts the revenues stemming from the optimal information structure. Clearly the revenues increase with the number of bidders and approach the feasible maximum which is represented by the expected social welfare. In comparison, the dashed line represents the revenues for the auctioneer when all bidders have perfect information. We simply refer to it as the Myerson auction.

Insert Figure 4 about here

Finally, Figure 4 depicts the relative revenue losses of the Myerson auction and the relative revenue losses of the optimal information structure, as the number of bidders increases. Relative revenue losses are calculated as the difference between the total surplus and the optimal auction divided by the social surplus. The revenue loss for the optimal information structure increase initially and then decrease, as the number of bidders increase. Revenue losses under the Myerson auction always decrease, as the number of bidders increases. The optimal information structure extracts considerably more surplus than the Myerson auction.

4 Discrete Information Structures

We begin our formal analysis with a local argument regarding the suboptimality of information structures which allow for an uncountable number of signals and hence uncountable number of expected valuations. The argument is local in two respects. First, the argument proceeds by showing that the revenue of the auctioneer can always be increased by replacing an uncountable information structure of agent i on a small interval of realizations by a single realization. Second, the argument considers only the revenue resulting from a single bidder and leaves the allocations of the remaining bidders unchanged when modifying the information structure of agent i. The local comparison involves the incentive compatible revenues for the auctioneer when the distribution $G_i(w_i)$ is composed of an absolute continuous and discrete function. As the received results in the mechanism design literature consider either exclusively continuous or discrete distributions, we first state a general result about the revenues of the auctioneer in the presence of both types.

4.1 Incentive Compatible Revenues

In contrast to the standard optimal mechanism literature, we consider an optimal mechanism where a type can either have zero density, positive density or positive probability. At this stage we are merely interested in characterizing the expected revenues of the auctioneer from bidder i. For a given distribution of (expected) valuations denote by $\mathcal{M}_i$ the (countable) set of mass points under the distribution, $\mathcal{M}_i \triangleq \left\{w_i^1, ..., w_i^k, ...\right\}$, such that for every $w_i^k \in \mathcal{M}_i$,

$$g_i^k \triangleq G_i\left(w_i^k\right) - \lim_{w_i \uparrow w_i^k} G_i(w_i) > 0,$$

and we denote the probability of a mass point w_i^k by g_i^k. Next let $\mathcal{P}_i$ be the set which has positive density, or

$$\mathcal{P}_i \triangleq \{w_i \,|g_i(w_i) > 0\}.$$

Finally, denote by $\mathcal{O}_i$ the maximal union of open intervals $\left(w_i^{l^+}, w_i^{l+1^-}\right)$, or

$$\mathcal{O}_i \triangleq \bigcup_{l=1}^{\infty} \left(w_i^{l^+}, w_i^{l+1^-}\right),$$

such that for all $w_i \in \left(w_i^{l^+}, w_i^{l+1^-}\right)$ we have $g_i(w_i) = 0$. The set $\mathcal{O}_i$ is maximal if there is no other open set, say (w_i', w_i''), outside of $\mathcal{O}_i$ such that for all $w_i \in (w_i', w_i'')$, $g_i(w_i) = 0$. We can now describe the revenues the auctioneer receives from bidder i with a given prior distribution $G_i(w_i)$ and a given expected probability of winning $Q_i(w_i)$.

Theorem 1 (Revenues)
The expected revenues from bidder i are given by $R_i(G_i, Q_i)$, *with*

$$\begin{aligned} R_i(G_i, Q_i) &= \int_{\mathcal{P}_i} Q_i(w_i)\left(w_i - \frac{1 - G_i(w_i)}{g_i(w_i)}\right) g_i(w_i)\, dw_i \\ &\quad + \sum_{k=1}^{\infty} Q_i\left(w_i^k\right) w_i^k g_i^k \\ &\quad - \sum_{l=1}^{\infty} Q_i\left(w_i^{l+}\right)\left(w_i^{l+1^-} - w_i^{l^+}\right)\left(1 - G_i\left(w_i^{l^+}\right)\right) \end{aligned} \tag{2}$$

subject to $Q_i(\cdot)$ *being non-decreasing.*

Proof. See appendix. ■

We remark that the probability $Q_i\left(w_i^{l^+}\right)$ which appears in the final sum of expression (2) is arbitrary in the sense that it could be replaced with any probability $Q_i(w_i)$ subject to $w_i \in (w_i^{l+}, w_i^{l+1^-})$ and the mechanism would still implement the same allocation of the object. The probability $Q_i\left(w_i^{l^+}\right)$ is the probability which maximizes the revenues of the auctioneer for any given rule of probabilistic allocations $\{Q_i\}_{i\in\mathcal{I}}$.

The revenue formula (2) becomes more transparent in some special cases of interest. First, suppose that the distribution function has strictly positive density everywhere and no mass points. Then the revenues can be written simply as:

$$\int_0^1 Q_i(w_i)\left(w_i - \frac{1 - G_i(w_i)}{g_i(w_i)}\right) g_i(w_i)\, dw_i,$$

where the central element is given by the familiar expression of the virtual utility:

$$w_i - \frac{1 - G_i(w_i)}{g_i(w_i)}.$$

The second special case of interest is the case when the distribution function has zero density everywhere and contains only mass points. In this case, $w_i^k = w_i^{l^-} = w_i^{l^+}$ for all $k = l$. Then the revenue for the auctioneer can be written as the sum:

$$\sum_{k=1}^{\infty} Q_i\left(w_i^k\right)\left(w_i^k - \left(w_i^{k+1} - w_i^k\right)\frac{1 - G_i\left(w_i^k\right)}{g_i^k}\right) g_i^k.$$

The similarity with the case of positive density is immediate. The modification due to the discreteness appears in the obvious places. The density $g_i(w_i)$ is now replaced by the positive probability g_i^k and the local change $dw = 1$ is being replaced by the discrete change between w_i^k and w_i^{k+1}, or

$w_i^{k+1} - w_i^k$. Finally, consider the case where the density is positive everywhere, but the distribution has mass points as well. In this case, the third term in (2) vanishes as there are no open sets with zero density and the mass points appear without virtual costs. This formulation may appear to be in sharp difference with the positive density case, but if we consider a mass point to be the limit point in a distribution with positive density, then the apparent distinctiveness disappears. In the limit, the density of mass point becomes arbitrarily large and as $g_i(w) \to \infty$, the virtual utility converges towards the value at that point, when we hold the value of the distribution function at w_i fixed, or

$$w_i - \frac{1 - G_i(w_i)}{g_i(w_i)} \to w_i.$$

Thus, the behavior of the revenues at the mass point can easily be understood from the limiting behavior of a distribution with positive density exclusively.

4.2 Informational Bundling

The representation of the expected revenues given in Theorem 1 allows us to undertake the next step. We are going to consider the revenue the auctioneer receives from a bidder i on an arbitrarily small interval of valuations in which the distribution function has positive density. We argue that the revenue resulting from this interval is dominated by the revenue generated if the interval with positive density is replaced with a single mass point. The revenue characterization in Theorem 1 allow us to perform this comparison for a single bidder and independent of the other bidders. We consider the behavior of the revenues on a small interval $[z, z + \varepsilon]$ by studying the marginal behavior of the revenues, as ε is increased. We find that the marginal return at $\varepsilon = 0$ are positive, as expected, and identical for the interval and its replacement in form of a mass point. However, when we look at the second derivative, we shall observe that revenues rise faster for the mass point than the equivalent interval. Essentially, this argument allows us to conclude that the distribution function $G_i(w_i)$ can be suitable modified by replacing intervals with mass points, whenever such intervals exist. Every such operation will lead to higher revenues. Thus, we show that the optimal information structure has to be discrete.

Theorem 2 (Discreteness)
Every (non-discrete) information structure is strictly dominated by a discrete information structure.

Proof. See appendix. ■

As the central argument of the proof might be lost in the details, we try to give a brief account of its structure. Consider an interval $[z, z + \varepsilon] \subset [0, 1]$ on which the distribution function has positive density, or $g_i(w_i) \geq 0$. The revenues from this interval for the auctioneer are given by:

$$R_z(\varepsilon) \triangleq \int_z^{z+\varepsilon} Q_i(w_i)\left(w_i - \frac{1-G_i(w_i)}{g_i(w_i)}\right) g_i(w_i)\, dw_i. \tag{3}$$

Suppose we replace the interval of valuations with a single mass point. The probability of the mass point is given by

$$g_z(\varepsilon) \triangleq \int_z^{z+\varepsilon} g_i(w_i)\, dw_i.$$

The associated conditional expected value on the mass point is given by

$$w_z(\varepsilon) \triangleq \frac{\int_z^{z+\varepsilon} w_i g_i(w_i)\, dw_i}{\int_z^{z+\varepsilon} g_i(w_i)\, dw_i}, \tag{4}$$

and the conditional expected probability of winning is

$$Q_z(\varepsilon) \triangleq \frac{\int_z^{z+\varepsilon} Q_i(w_i)\, g_i(w_i)\, dw_i}{\int_z^{z+\varepsilon} g_i(w_i)\, dw_i}. \tag{5}$$

As we remove the density on the interval $[z, z+\varepsilon]$ and replace it by a single probability mass at $w_z(\varepsilon)$, we change the revenues locally in two ways. First, we introduce a mass point where there was none before and, second, we introduce two intervals with zero density. We may use Theorem 1 to evaluate how these changes affect the revenues. The new revenues on the interval are:

$$\begin{aligned} \widehat{R}_z(\varepsilon) \quad &\triangleq \quad w_z(\varepsilon)\, p_z(\varepsilon)\, Q_z(\varepsilon) - (w_z(\varepsilon) - z)\, Q_i(z)\,(1 - G_i(z)) \\ &\quad - (z + \varepsilon - w_z(\varepsilon))\, Q_z(\varepsilon)\,(1 - G_i(z+\varepsilon)). \end{aligned} \tag{6}$$

As we transform the interval into a single mass point, we maintain the expected probability by which we assign the object to bidder i. While the expected probability $Q_z(\varepsilon)$ may not necessarily constitute the optimal solution from the auctioneer's point of view, it implies that the incentives for all bidders except i remain unmodified and hence we concentrate without loss of generality on the revenues resulting from bidder i. The proof then proceeds to show that $0 < R'_z(0) = \widehat{R}'_z(0)$, while $0 < R''_z(0) < \widehat{R}''_z(0)$.

As the marginal revenues increases faster for the mass point than for the interval, we conclude that any interval with positive density when replaced by a mass point yields higher revenues for the auctioneer. The basic trade-off can be traced back to the local revenues in (3) and (6) and can be conceived as a trade-off between (virtual) efficiency and information rent. By bundling all realizations in the interval $[z, z+\varepsilon]$ the auctioneer looses the ability to vary the probability of winning as a function of the realization in the interval $[z, z+\varepsilon]$. If the auctioneer were to maximize the social utility, then this would translate into a social loss. As he only maximizes the virtual utility, it represents a loss in virtual utility. Yet, by changing the support of the distribution function, the auctioneer actually changes the virtual utility. The change in the virtual utility arises as the continuum of incentive constraints on the interval are replaced by just two (downward) incentive constraints: (i) from $w_z(\varepsilon)$ to z and (ii) from $z+\varepsilon$ to $w_\varepsilon(z)$. In fact, the difference between $R_z''(0)$ and $\widehat{R}_z''(0)$ can be traced to the fact that the initial loss in efficiency is dominated by the gain in surplus due to a depression of the information rent.

Finally, we might ask whether these modifications in the distribution function $G_i(w_i)$ can be supported and generated by an appropriate information structure, which is the primitive of our model. As every expected value w with positive density $g_i(w_i) > 0$ must be generated by at least one signal s_i, let

$$w_i^{-1}(A_i) \triangleq \{s_i \,|w_i(s_i) \in A_i\}.$$

Since two signals s_i and s_i' may generate the same expected value, or $w_i(s_i) = w_i(s_i')$, the mapping $w_i^{-1}(\cdot)$ may be a correspondence rather than a function. Suppose the set A_i is given by $A_i = [z, z+\varepsilon]$. The replacement of the interval by a single mass point in the distribution function can now be easily mirrored by changes in the information structure. Consider the signals defined by $w_i^{-1}(A_i)$. By replacing the entire set $w_i^{-1}(A_i)$ by a single signal $s_i = s_{A_i}$ and transferring the original probability of the set A_i onto the signal s_{A_i}, we change the information structure such that the resulting $G_i(w_i)$ is precisely the modified distribution used in Theorem 2. As every mass point in the distribution $G_i(w_i)$ can be generated without loss of generality by a single signal s_i, it follows that a discrete distribution function $G_i(w_i)$ can always be supported by a countable set S_i of signals. Thus, we can translate the discreteness in the distribution function into a discrete information structure.

5 Optimal Information Structures

In this section we prove that the optimal information structure can be represented as a finite partition for every bidder i. By the results of the preceding section, we know that the information structure has to be discrete. Therefore, Subsection 5.1 briefly states some results regarding the optimal auction with a countable number of types each having positive probability. Subsection 5.2 uses the revenue structure to show that the optimal information structure has to be a countable and monotone partition. As a by-product of the proofs we obtain several interesting results about the virtual utilities and the structure of the optimal auction. Subsection 5.3 refines the results to conclude that the partition for each bidder has to be finite.

5.1 Optimal Auction with Discrete Types

This section characterizes the optimal auction when the distribution of valuations for each bidder is given by an arbitrary probability distribution function G with countable support. We recall that the valuations at these points are denoted by w_i^k and without loss of generality:

$$0 \leq w_i^1 < w_i^2 < .... < w_i^k < .... \leq 1.$$

The probability of the realization w_i^k is denoted as before by g_i^k, and let $G_i^k \triangleq G_i\left(w_i^k\right)$ and similarly, $Q_i^k \triangleq Q_i\left(w_i^k\right)$.

Theorem 1 characterizes the revenues of the auctioneer from bidder i as a function of the expected probability of winning $Q_i\left(w_i\right)$ with a value w_i. The interaction with the valuation of the other bidders in the earlier formulation was represented by expectations over the valuations w_{-i}. Now, we disaggregate the expression and consider the dependence on the realizations of all valuations explicitly. The revenue of the auctioneer from all bidders is given by:

$$R\left(G,q\right) \triangleq \sum_{k_1=1}^{\infty} \cdots \sum_{k_I=1}^{\infty} \left[\sum_{i=1}^{I} q_i\left(w_1^{k_1},...,w_I^{k_I}\right)\left[w_i^{k_i} - \left(w_i^{k_i+1} - w_i^{k_i}\right)\frac{1-G_i^{k_i}}{g_i^{k_i}}\right]\prod_{i=1}^{I} g_i^{k_i}\right], \tag{7}$$

where $q_i\left(w\right) \geq 0$ and $\sum_{i=1}^{I} q_i\left(w\right) \leq 1$. The formula (7) is based on the general expression for the revenues in Theorem 1. Define the virtual utility with discrete types by:

$$\gamma_i^k \triangleq w_i^k - \left(w_i^{k+1} - w_i^k\right)\frac{1-G_i^k}{g_i^k}.$$

As in Myerson (1981), the optimal allocation policy $q^*\left(w\right) = \left(q_1^*\left(w\right),...,q_I^*\left(w\right)\right)$ can be determined by pointwise optimization provided that the virtual utility γ_i^k is non-decreasing in k for every agent i. If the virtual utilities γ_i^k for a given distribution G_i^k were not monotone, then the optimal auction would be subject to a similar "ironing out" procedure as necessary in an optimal auction with a continuum of types. The basic element in the former procedure is to maintain the expected probability Q_i^k constant over a set of types which covers the non-monotonicity. As the constant probability essentially implies that the incentives and revenues are constant on it, the question arises as to whether the auctioneer has any interest in distinguishing between different types in this set. In fact, as the information structure is chosen by the auctioneer, he may wish to bundle types to which identical allocations have to be offered in any case. In other words, when the auctioneer can choose the information structure for the bidders, the "ironing out" of non-monotonicities in the virtual utility may be achieved by a sufficient coarsening of the information structure rather than through constant winning probabilities of the form: $Q_i^k = Q_i^{k+1}$. The consequence of this argument leads to the next result.

Theorem 3 (Monotone Virtual Utilities)

The optimal virtual utilities are strictly increasing.

Proof. See appendix. ■

We can now readily describe the optimal auction mechanism when the virtual utilities are monotone. The characterization is the exact discrete type analog to the celebrated optimal auction result for 'regular environments' by Myerson (1981) with a continuum of types.

Corollary 1

Suppose the virtual utilities are increasing for every agent. The optimal auction is described by:

1. $q_i\left(w_1^{k_1}, ..., w_I^{k_I}\right) > 0 \Rightarrow \gamma_i^{k_i} \geq 0 \wedge \gamma_i^{k_i} \geq \gamma_j^{k_j}, \forall j$;
2. $\max\left\{\gamma_1^{k_1}, ..., \gamma_I^{k_I}\right\} > 0 \Rightarrow \sum_{i=1}^{I} q_i\left(w_1^{k_1}, ..., w_I^{k_I}\right) = 1$.

Proof. See appendix. ■

By Theorem 3, we can describe the set of virtual utilities for bidder i by an ordered set $\Gamma_i = \left\{\gamma_i^1, ..., \gamma_i^k, ...\right\}$, with $\gamma_i^1 < \gamma_i^2 < ... < \gamma_i^k < ....$.

As we argued earlier that additional valuations are associated with additional incentive constraints, the optimal information structure strikes a balance between allocational efficiency and informational rent. This has some immediate implications for the structure of the set of virtual utilities Γ_i. Consider two adjacent and positive virtual utilities by agent i, say γ_i^k and γ_i^{k+1}. According to Corollary 1 they will receive different winning probabilities if and only if they bracket the virtual utility of some other agents. By contrast, γ_i^k and γ_i^{k+1} receive the same winning probability if they do not bracket any other virtual utility realization. In the later case, the realizations w_i^k and w_i^{k+1} have to satisfy the same incentive constraints and assignment rule. Thus, they can be joined to a single realization with an expected value of

$$\bar{w}_i^k \triangleq \frac{w_i^k g_i^k + w_i^{k+1} g_i^{k+1}}{g_i^k + g_i^{k+1}}.$$

This replacement does not reduce (virtual) efficiency. However, revenues increase due to lower information rents associated with fewer incentive constraints.

Theorem 4 (Adjacent and Asymmetric Virtual Utilities)

1. *For every* γ_i^k: $\left\{\gamma_j^k \,\middle|\, \gamma_i^k < \gamma_j^k < \gamma_i^{k+1}\right\} \neq \emptyset$.
2. $\exists i, j$ *such that* $\Gamma_i \neq \Gamma_j$.

Proof. See appendix. ■

A direct consequence of the alternating structure of the virtual utilities is the asymmetry of the virtual utilities indicated by the second part of Theorem 4. With two bidders, the same argument leads immediately to a stronger result, namely that $\Gamma_i \cap \Gamma_j = \emptyset$. With more than two bidders, our argument does not preclude the possibility that some bidders may have virtual utilities in common. We conjecture that this will not occur under an optimal information structure.

With Theorem 4 in place we may suppose without loss of generality that the virtual utilities across bidders are ordered as follows:

$$\gamma_1^1 \leq \gamma_2^1 \leq ... \leq \gamma_I^1 < \gamma_1^2 \leq \gamma_2^2 \leq .... \leq \gamma_{I-1}^k \leq \gamma_I^k < \gamma_1^{k+1} \leq ... \tag{8}$$

While some virtual utilities in (8) may not be generated by the optimal information structure, we can always pretend they existed nonetheless and set $g_i^k = 0$. We can now write the revenues of an optimal auction, where the virtual utilities generated by the distribution functions $G = \{G_i\}_{i=1}^I$ satisfy the properties derived in Theorem 3 and 4, as follows:

$$R(G) = \sum_{i=1}^{I} \sum_{k=1}^{\infty} \left(\max\left\{\gamma_i^k, 0\right\} g_i^k \prod_{j \neq i} G_j^{k(i)} \right),$$

where $G_j^{k(i)}$ is defined by:

$$G_j^{k(i)} = \begin{cases} G_j^k, & \text{if} \quad j < i, \\ G_j^{k-1}, & \text{if} \quad j > i. \end{cases} \tag{9}$$

As the expected winning probabilities Q_i^k in the optimal auction are given by:

$$Q_i^k \triangleq \prod_{j \neq i} G_j^{k(i)},$$

we can write the revenues to be:

$$R(G) = \sum_{i=1}^{I} \sum_{k=1}^{\infty} \left(\max\left\{\gamma_i^k, 0\right\} g_i^k Q_i^k \right). \tag{10}$$

5.2 Monotone Partitions

The local analysis in the previous section allowed us to conclude that the optimal information structure will have a countable signal space S_i. The countable support result for the marginal distribution of the signals however doesn't permit any further inferences about the joint distribution or the conditional distribution of signals. Now, we extend the local to a global analysis taking into account explicitly the interaction among bidders. We have shown earlier that the optimal information structure has to be contained in the class of discrete information structures. Thus, the global analysis can be performed directly on the optimal auction for discrete types which we characterized above.

A partitional information structure can be represented without recourse to a joint distribution over the space of valuations and signals by a partition of the original space V_i. A partition is a collection of subsets $\mathcal{S}_i = \left\{S_i^k\right\}$ such that for all k, k' we have $S_i^{k'} \cap S_i^k = \emptyset$ and

$$\bigcup_{k=1}^{\infty} S_i^k = V_i.$$

The properties of monotonicity and finiteness can then be simply be restated on the space of valuations. The partition is *monotone* if for any $v_i, v_i' \in S_i^k$, $\lambda v_i + (1-\lambda) v_i' \in S_i^k$ for all $\lambda \in [0,1]$. The partition is finite if the collection of subsets has a finite cardinality.

Theorem 5 (Monotone Partition)

1. *The optimal information structure is a partition.*

2. *The optimal partition is monotone.*

Proof. See appendix. ■

The monotonicity of the partition as well as the optimality of the partition itself result from the same basic argument. In fact, the partitional result can be viewed as a mere complication arising from additional noise in the information structure. Therefore, we content ourselves to give a brief outline of the proof of the monotonicity property here. Since, we can restrict our attention from now on to partitions of the space of valuations, we can state all probabilities in terms of the original distributions $F_i(v_i)$. Therefore, suppose that the information structure can be described by a partition $\mathcal{S}_i$ for every bidder i. The probability f_i^k of the event $S_i^k \in \mathcal{S}_i$ is given by

$$f_i^k \triangleq \int_{v_i \in S_i^k} f_i(v_i)\, dv_i.$$

The realization of the event S_i^k leads agent i to adopt a conditional expectation

$$w_i^k \triangleq \frac{\int_{v_i \in S_i^k} v_i f_i(v_i)\, dv_i}{\int_{v_i \in S_i^k} f_i(v_i)\, dv_i}.$$

A point z is called a partitional point of S_i^k and S_i^l if for every neighborhood of z, or for all $\varepsilon > 0$ and $B_\varepsilon(z)$, there exists $v_i, v_i' \in B_\varepsilon(z)$ such that $v_i \in S_i^k$ and $v_i' \in S_i^l$. We denote by $\mathbf{z}_i$ the vector of all *partitional points* between any two (or multiple) partitions of bidder i and by $\mathbf{z} = (\mathbf{z}_1, ..., \mathbf{z}_I)$ the vector of all partitional points. The revenues of the auctioneer in the optimal auction can be described as a function of the partitional points $\mathbf{z}$, or $R(\mathbf{z})$. The idea of the proof can now be described as follows. Consider an arbitrary point v_i which may be either allocated to the partitional element S_k^i or S_l^i. The difference in the marginal revenue of this assignment to either of the partitional elements k and l can be written as follows:

$$a_i^{k,l} + \left(b_i^l - b_i^k\right) v_i,$$

where the parameters $a_i^{k,l}, b_i^k$ and b_i^l are determined by the size and winning probabilities of the given partitional elements S_k^i and S_l^i. It follows that every partitional point z_i between S_i^k and S_i^l must satisfy at the optimum

$$\frac{\partial R(\mathbf{z})}{\partial z_i} = a_i^{k,l} + \left(b_i^l - b_i^k\right) z_i = 0. \tag{11}$$

In other words, the first order conditions can be described at the optimum as linear functions of the partitional points, where the values of the parameters are given by the partitions which are separated by the point z_i. By the strict single crossing property of the linear functions, it follows that every two elements of the partition can have at most one partitional point in common (provided that $a_i^{k,l} = b_i^l = b_i^k = 0$ can be excluded from our considerations). Moreover, as the slope parameters in (11) depend additively on the elements S_i^k and S_i^l, it follows from the linearity of the first order conditions that only adjacent partitional elements can have partitional points, or $l = k + 1$. This argument establishes the monotonicity of the partition.

The partitional property of the optimal information structure is reminiscent of the partitional structure of the strategies in the cheap talk games investigated by Crawford & Sobel (1982). While the partitions are endogenous and lead to coarse decision making in both settings, the reasons why partitions emerge are very different. In Crawford & Sobel (1982) the principal cannot offer monetary transfers to align the incentives of the agent with his own. The partial alignment can only arise through the actions which are contingent on the message sent by the agent. By allowing the messages to be sufficiently coarse, both, agent and principal, prefer the action chosen based on limited information to any of the other, few and sufficiently distinct, equilibrium actions. In our setting the principal can use monetary transfers to align the preferences of the agent. However, the arguments showed that it is too costly for him to obtain alignment for all possible signals, and he prefers to adopt a coarser decision rule to limit the costs of aligning the incentives. The principal achieves this formally by allowing fewer signals and, therefore, fewer incentive constraints.

5.3 Finiteness

We may summarize the results obtained up to this point as follows. Theorem 2 shows that every continuous information structure is dominated by a discrete information structure. In addition, Theorem 5 shows that the information structure has the form of a monotone partition. We further demonstrated that the induced virtual utilities are monotone for each bidder (Theorem 3) and asymmetric in the sense that the virtual utilities are distinct across bidders (Theorem 4). What is yet missing in the description of the optimal information structure is the determination as to whether an optimal information structure *exists* and whether it is finite or rather infinite.

By Theorem 5, we can describe the set Z_i of partitional points of bidder i by:

$$Z_i = \left\{0, z_i^1, z_i^2, ..., z_i^k, ...\right\}$$

and an element S_i^k of the partition as an interval $S_i^k = [z_i^{k-1}, z_i^k)$. We start the remaining investigation by considering finite partitions. Denote by K the maximal number of elements in the partition of every agent i.

Theorem 6 (Finite Partitions) *For every fixed $K < \infty$, an optimal partition exists.*

Proof. See appendix. ■

Theorem 6 establishes the existence of an optimal partition if we restrict the maximal number of elements in every partition to be K. Next, we consider properties of the optimal partition as a function of K. Suppose for an increasing sequence of K, the restriction on K would remain a binding constraint in the sense that at least for some agent i, the optimal partition contains exactly K elements. It would then follow that there are at least two adjoining intervals in the partition of an agent i which will become arbitrarily small as K becomes arbitrarily large. We investigate the implication for the first order conditions and the associated expected winning probabilities of agent i, if two adjoining intervals become arbitrarily small.

Proposition 1 (Winning Probabilities)
For every $\varepsilon > 0$, there exists $\delta(\varepsilon) > 0$ such that for all i and k:

$$z_i^k - z_i^{k-1} \leq \delta(\varepsilon) \text{ and } z_i^{k+1} - z_i^k \leq \delta(\varepsilon),$$

implies

$$\left| \frac{Q_i^k - Q_i^{k-1}}{Q_i^{k+1} - Q_i^k} - 1 \right| \leq \varepsilon.$$

Proof. See appendix. ■

Proposition 1 allows us to conclude that as long as the intervals become sufficiently small, the winning probability Q_i^k has to increase proportionally, independently of the relative size of the intervals $[z_i^{k-1}, z_i^k)$ and $[z_i^k, z_i^{k+1})$. This property plays an essential role in the next argument, in which we extend the earlier result on the optimality of discrete versus continuous information structures (Theorem 2) to show that the discrete information structure cannot be too fine either. Suppose we join two adjacent intervals to form a single interval. The expected value on the interval is now given by:

$$\bar{w}_i^k \triangleq \frac{w_i^k f_i^k + w_i^{k+1} f_i^{k+1}}{f_i^k + f_i^{k+1}}.$$

We compare the revenue for the auctioneer under the original and the modified partition. The comparison proceeds by maintaining the winning probabilities Q_i^k of the original partition even when computing the revenues associated with the modified partition. This allows us to compare the revenue across the two partitions by simply analyzing the revenue resulting from the partitional elements $[z_i^{k-2}, z_i^{k-1}), [z_i^{k-1}, z_i^k)$ and $[z_i^k, z_i^{k+1})$. Correspondingly, we denote the revenue difference between the modified and the original segment k by D_i^k. The modified partition displays different virtual utilities, denoted by $\bar{\gamma}_i^k$, where the modification arises locally due to the combination of w_i^k and w_i^{k+1}, namely

$$\bar{\gamma}_i^{k-1} \triangleq w_{k-1}^i - \left(\bar{w}_i^k - w_i^{k-1}\right) \frac{1 - F_i^{k-1}}{f_i^{k-1}},$$

and

$$\bar{\gamma}_i^k = \bar{\gamma}_i^{k+1} \triangleq \bar{w}_i^k - \left(w_i^{k+2} - \bar{w}_i^k\right) \frac{1 - F_i^{k+1}}{f_i^k + f_i^{k+1}}.$$

The difference in revenues in segment k is then defined formally by $D_i^k \triangleq \left(\bar{\gamma}_i^k - \gamma_i^k\right) f_i^k$.

Proposition 2 (Pairwise Coarsening)

There exists $\bar{\varepsilon} > 0$ such that for all $\varepsilon \in (0, \bar{\varepsilon})$, for all i and k,

$$z_i^k - z_i^{k-1} \leq \varepsilon \text{ and } z_i^{k+1} - z_i^k \leq \varepsilon,$$

implies

$$D_i^{k-1} Q_i^{k-1} + D_i^k Q_i^k + D_i^{k+1} Q_i^{k+1} > 0.$$

Proof. See appendix. ■

The proof of this result is substantially more involved than the earlier result regarding the suboptimality of continuous information structures. The cause of the complication is that it is not sufficient anymore to look at small intervals for increasing, but otherwise arbitrary, winning probabilities Q_i^k. In fact, without the equilibrium bounds on the winning probabilities established in Proposition 1, the result is false if the intervals were to be decreasing too rapidly in size.

Theorem 7 (Finite Partition)
The optimal information structure exists and is a finite and monotone partition.

Proof. See appendix. ■

Theorem 7 shows that the optimal information structure in the class of discrete information structures is finite (and can be represented by a monotone partition). The proof of Theorem 7 could be easily extended to encompass uncountable and in particular a continuum of signals. Since the proof doesn't rely on any continuity or differentiability properties of the density $g_i(w_i)$, it shows the dispensability of the regularity assumption made earlier to permit a 'local' proof of the discreteness of the information structure.

At this point, we might recall the earlier example with a uniform density and the resulting optimal partition. The structure of the partition in the example may illustrate that the finiteness result is rather strong, as the cardinality of the partition is very small and increases only slowly in the number of bidders. Yet, we may expect that, as the number of bidders increases, the information rents become less important in the calculus of the seller and the incentives to compress information into partitions becomes less influential in the design problem. Therefore, we complete the analysis by a limiting results on the number of bidders.

6 Many Bidders and Sequential Sale

In this section we derive some further implications for the auction mechanism under the optimal information structure. Subsection 6.1 shows that, as the number of bidders participating in the auction increases, many bidder will be assigned arbitrarily fine partitions. Subsection 6.2 shows that the revenue of the optimal direct mechanism can be achieved by a sequential mechanism in which the seller makes a series of take-it-or-leave-it offers.

6.1 Many Bidders

Next, we show that, as the number of bidders increases, the information structure becomes finer. Theorem 7 establishes that information structures consist of finitely many intervals. Now, we assume that intervals cannot be smaller than of length ε. We show that eventually, as the number of agents increases, there are agents that adopt the finest information structure. We assume that interval points have to be multiples of ε, that is, $z_i^k = k \cdot \varepsilon$. Furthermore, an ε-partition denotes the partition $\{[(k-1) \cdot \varepsilon, k \cdot \varepsilon)\}_{k=1}^{K}$ with $K = \frac{1}{\varepsilon}$. It is a partition that consists of intervals of length ε.

Theorem 8

For any $\varepsilon > 0$ there exists a subset of bidders $I_1 \subset I$ such that as, $I \to \infty$, the subsequence, $I_1 \to \infty$, and all $i \in I_1$ have an ε-partition.

Proof. See appendix. ■

The theorem considers the calculus of allocative efficiency versus informational rents as the number of bidders increases. The theorem shows that as the number of bidder increases, the motive of allocative efficiency eventually comes to dominate the suppression of the informational rents with coarse information. With many bidders, competitive forces eliminate the informational rents of the bidders anyhow and, thus, the need to reduce the informational rents through coarse private information disappears. The proof assumes that the bidders are symmetric and have the same prior distribution F. This restriction is made to simplify the exposition of the proof and is not required. The proof could be easily extended to cover the case of replicating a given set of bidders, with possible asymmetries among the initial set of bidders.

The proof proceeds by induction. The main idea of the proof is the following: First, we show that there exists a subsequence of bidders that have the interval $[1-\varepsilon, 1]$ in their partition. We obtain this property by establishing a contradiction. If the property holds, then the revenues approach $1 - \frac{\varepsilon}{2}$, as the number of bidders increases, since the probability that at least on bidder will obtain a valuation contained in the highest interval approaches one. Suppose that the property does not hold and at most finitely many bidders have a partition on the top. The highest valuation of the remaining bidders equals at most $1 - \varepsilon$ and, therefore, there exists a $\delta > 0$ such that revenues are less than $1 - \frac{\varepsilon}{2} - \delta$. Now, eventually, as the number of bidders increases, the revenues are dominated by the scheme that approaches revenues of $1 - \frac{\varepsilon}{2}$. This contradiction establishes the first part.

The induction argument proceeds by looking at lower intervals and uses our earlier results on the structure of the optimal revenues. These results permit us to decompose the optimal revenues into revenues obtained from virtual valuations exceeding a certain threshold and revenues below the threshold. In order to study the seller's maximized revenues, we can look at the revenues below the threshold independent of the revenues above the threshold. We show that there must be a subsequence of bidders that have the highest possible virtual valuation conditional on virtual valuations being below the threshold. The argument is similar to the above described argument and we omit it in this discussion. Since, the number of intervals in the partition are finite, we have shown that there must be a subsequence of bidders $I_1 \to \infty$ each having an ε-partition, as the number of bidders gets large, $I \to \infty$.

6.2 Sequential Sale

A remarkable feature of the optimal information structure is the fact that the set of virtual utilities is distinct for every agent i. As shown in Theorem 4, the asymmetry in the virtual utilities extends to environment where the bidders, which are characterized by their prior distribution $F_i(v_i)$, are initial symmetric, or $F_i(v_i) = F(v_i)$ for all i. Moreover, the set of virtual utilities is formed such that two adjacent virtual utilities are never generated by the same bidder. Therefore, the finiteness and asymmetry of the virtual utilities permit a very simple implementation of the optimal auction in the form of a sequential posted price mechanism rather than simultaneous bidding mechanism. Given the information structure, the auctioneer can device a sequential mechanism in which he successively asks individual bidders whether they would like to receive the object at a fixed price.

Theorem 9
The revenue of optimal direct revelation mechanism can be realized by the following sequential and indirect mechanism: The auctioneer makes exclusive take-it-or-leave-it offers, in which no bidder receives two subsequent offers.

Proof. See appendix. ■

In the current setting, the sequential selling mechanism only allows the auctioneer to achieve the same expected revenue as the optimal simultaneous auction. However, if we were to add costs to the provision and/or assignment of information structures, then the sequential mechanism would come to strictly dominate the simultaneous mechanism as it would economize on the cost of information.

7 Conclusion

This paper reconsidered the design of the optimal auction by making the information structure an integral part of the design problem. Notable features of the optimal information structure were the partitional character, the finiteness of the partition and, therefore, private types, and the asymmetry of the information structures.

While the characterization of the qualitative features of the optimal information structure was reasonably complete, several additional results might be obtainable. First, it would seem that the fineness of the partition should be a monotone function of the number of participating bidders. Second, it would be interesting to obtain a rate of convergence result for the optimal information structure towards the perfect information structure as the number of bidders increases. Third, it would be interesting to know how fast the revenues of the auctioneer converge to full surplus extraction with the optimal information structure vis-a-vis the perfect information structure. The opening example with the uniform density suggests that these questions are well defined and might allow for clear answers, but we have not yet pursued these questions in sufficient detail.

The basic allocation problem considered in this paper is the assignment of a single unit to one among many bidders. The complete characterization of the optimal mechanism for any arbitrary information structure facilitated the task of finding the optimal information structure. However, the basic arguments and results seem to be robust to general mechanism design problems. Therefore, it may be possible to generalize the results here beyond the auction environment. Adverse selection problems with a single agent, as in regulation and procurement environments, may present more difficult allocation problems, due to nonlinearity of the optimal solution, but at the same time, the endogeneity of the information structure may particularly relevant for contracting purposes in this environment.

Finally, we observe that our results do not bear directly on the linkage principle (Milgrom & Weber (1982)). In our set-up, the auctioneer influences the informativeness of every signal received by the bidders, yet he never observe the signal realization, which remains private information for the agent. Thus, if the auctioneer provides a bidder with a more informative signal, he increases the amount of private information to which the agent has access to, and thus the informational rent of the bidder.

8 Appendix

The appendix contains the proofs to all proposition and theorem in the text as well as some auxiliary results and proofs.

Proof of Theorem 1. We consider the revenue resulting from a single bidder i and therefore restrict attention to the elements in the mechanism which affect bidder i. It is well known, see Mas-Collel, Whinston & Green (1995), Proposition 23.D.2, that the allocation $\{Q_i(w_i), T_i(w_i)\}$ is Bayesian incentive compatible if and only if:

1. $Q_i(w_i)$ is nondecreasing,

2. the equilibrium utility satisfies:

$$U_i(w_i) = U_i(0) + \int_0^{w_i} Q_i(u_i)\, du_i.$$

Notice, that the equilibrium utility of the agent is increasing in $Q_i(u_i)$. Next, consider the set $\mathcal{O}_i$ of open intervals which have zero density everywhere. The probability $Q_i(w_i)$ with which bidder i receives the object for all $w_i \in \mathcal{O}_i$ affects the utility of the bidder (and, thus, the revenues to the auctioneer) only through its appearance in the equilibrium utility of the agent. Clearly, in order to maximize revenues the auctioneer would like to minimize $Q_i(w_i)$ for all $w_i \in \mathcal{O}_i$. Since $Q_i(w_i)$ is restricted to be nondecreasing, we can set $Q_i(w_i) = Q_i\left(w_i^{l+}\right)$ for $w_i \in (w_i^{l+}, w_i^{l+1^-})$. We can write the equilibrium utility of the agent for $w_i \notin \mathcal{O}_i$ as

$$U_i(w_i) = U_i(0) + \sum_{l=1}^{l(w_i)} \int_{w_i^{l^-}}^{w_i^{l^+}} Q_i(u_i)\, du_i + \sum_{l=1}^{l(w_i)} Q_i\left(w_i^{l^+}\right)\left(w_i^{l+1^-} - w_i^{l^+}\right) + \int_{w^{l(w_i)+1^-}}^{w_i} Q_i(u_i)\, du_i,$$

where $l(w_i) = \max\left\{l \,\middle|\, w_i^{l^+} \leq w_i\right\}$ and $w_i^{1^-} = 0$. The revenues of the auctioneer can then be written as:

$$R_i(G_i, Q_i) = \int_0^1 \left[Q_i(w_i)\, w_i - U_i(w_i)\right] g_i(w_i)\, dw_i + \sum_{k=1}^{\infty} \left[Q_i\left(w_i^k\right) w_i^k - U_i\left(w_i^k\right)\right] g_i^k.$$

Integration by parts and the equivalent thereof for the discrete probabilities leads us to the formula in (2).■

The following lemma delivers auxiliary calculations for the proof of Theorem 2.

Lemma 1 *The conditional means $w_z(\varepsilon)$ and $Q_z(\varepsilon)$, as defined in (4) and (5), satisfy:*

$$w_z'(0) = \frac{1}{2},$$

and

$$Q_z'(0) = \frac{1}{2} Q_i'(z).$$

Proof. Using the definition of the conditional mean, we have

$$w_z'(\varepsilon) = \frac{(z+\varepsilon)\, g_i(z+\varepsilon) \int_z^{z+\varepsilon} g_i(w_i)\, dw_i - \int_z^{z+\varepsilon} w_i g_i(w_i)\, dw_i g_i(z+\varepsilon)}{\left(\int_z^{z+\varepsilon} g_i(w_i)\, dw_i\right)^2}.$$

As numerator and denominator converge to zero as $\varepsilon \to 0$, we can use l'Hopital to obtain the following expression:

$$\lim_{\varepsilon \to 0} w_z'(\varepsilon) = g_i(z+\varepsilon) \times \lim_{\varepsilon \to 0} \frac{\int_z^{z+\varepsilon} g_i(w_i)\, dw_i}{2\int_z^{z+\varepsilon} g_i(w_i)\, dw_i g_i(z+\varepsilon)} = \frac{1}{2}.$$

Consider next the derivative of $Q_z'(\varepsilon)$

$$Q_z'(\varepsilon) = \frac{Q_i(z+\varepsilon)\, g_i(z+\varepsilon) \int_z^{z+\varepsilon} g_i(w_i)\, dw_i - \int_z^{z+\varepsilon} Q_i(w_i)\, g_i(w_i)\, dw_i g_i(z+\varepsilon)}{\left(\int_z^{z+\varepsilon} g_i(w_i)\, dw_i\right)^2},$$

which we may rewrite to obtain:

$$\frac{(Q_i(z+\varepsilon) - Q_z(\varepsilon))\, g_i(z+\varepsilon)}{\int_z^{z+\varepsilon} g_i(w_i)\, dw_i}.$$

By l'Hopital we get for $\varepsilon = 0$, the following result:

$$Q_z'(0) = g_i(z+\varepsilon) \times \lim_{\varepsilon \to 0} \frac{Q_i'(z+\varepsilon) - Q_z'(\varepsilon)}{g_i(z+\varepsilon)},$$

which leads to

$$Q_z'(0) = Q_i'(z) - Q_z'(0),$$

or

$$Q_z'(0) = \frac{1}{2} Q_i'(z),$$

which completes the proof. The (almost everywhere) differentiability of $Q_i(z)$ is guaranteed by the differentiability of $g_i(w_i)$. ■

Proof of Theorem 2. The proof is by contradiction. Suppose therefore that the information structure has a positive density and no mass points over some interval $[z, z+\varepsilon]$ for some agent i. We show that if replace the random variable w_i on the interval $[z, z+\varepsilon]$ by the expected value of w_i conditional on $w_i \in [z, z+\varepsilon]$, then we can increase the payoff of the auctioneer. Differentiating $R_z(\varepsilon)$, as defined in (3), once we get:

$$R_z'(\varepsilon) = \left(z + \varepsilon - \frac{1 - G_i(z+\varepsilon)}{g_i(z+\varepsilon)}\right) g_i(z+\varepsilon)\, Q_i(z+\varepsilon),$$

and evaluated at $\varepsilon = 0$,

$$R'_z(0) = \left(z - \frac{1 - G_i(z)}{g_i(z)}\right) g_i(z) Q_i(z).$$

The first derivative of $\hat{R}_z(\varepsilon)$, where the later is defined in (6), is given by:

$$\begin{aligned}
\hat{R}'_z(\varepsilon) &= (w'_z(\varepsilon)(1 - G_i(z)) - (1 - G_i(z+\varepsilon)) + (z+\varepsilon) g_i(z+\varepsilon)) Q_z(\varepsilon) \\
&\quad + (w_z(\varepsilon)(1 - G_i(z)) - (z+\varepsilon)(1 - G_i(z+\varepsilon))) Q'_z(\varepsilon) \\
&\quad - w'_z(\varepsilon)((1 - G_i(z)) Q_i(z)),
\end{aligned}$$

and evaluated at $\varepsilon = 0$ after using Lemma 1:

$$\begin{aligned}
\hat{R}'_z(0) &= \left(\frac{1}{2}(1 - G_i(z)) - (1 - G_i(z)) + z g_i(z)\right) Q_i(z) \\
&\quad + (z(1 - G_i(z)) - z(1 - G_i(z))) \frac{1}{2} Q_i(z) \\
&\quad - \frac{1}{2}((1 - G_i(z)) Q_i(z)),
\end{aligned}$$

which after some cancellations is equivalent to $R'_z(0)$. Consider next the second derivative of $R_z(\varepsilon)$:

$$\begin{aligned}
R''_z(\varepsilon) &= (2g_i(z+\varepsilon) + (z+\varepsilon) g'_i(z+\varepsilon)) Q_i(z+\varepsilon) \\
&\quad + ((z+\varepsilon) g_i(z+\varepsilon) - 1 + G_i(z+\varepsilon)) Q'_i(z+\varepsilon),
\end{aligned}$$

and evaluating at $\varepsilon = 0$, we get

$$R''_z(0) = (2g_i(z) + z g'_i(z)) Q_i(z) + (z g_i(z) - (1 - G_i(z))) Q'_i(z).$$

In contrast, the second derivative of $\hat{R}_z(\varepsilon)$ is:

$$\begin{aligned}
\hat{R}''_z(\varepsilon) &= (w''_z(\varepsilon)(1 - G_i(z)) + 2g_i(z+\varepsilon) + (z+\varepsilon) g'_i(z+\varepsilon)) Q_z(\varepsilon) \\
&\quad + 2(w'_z(\varepsilon)(1 - G_i(z)) + (z+\varepsilon) g_i(z+\varepsilon) - (1 - G_i(z+\varepsilon))) Q'_z(\varepsilon) \\
&\quad + (w_z(\varepsilon)(1 - G_i(z)) - (z+\varepsilon)(1 - G_i(z+\varepsilon))) Q''_z(\varepsilon) \\
&\quad - w''_z(\varepsilon)(1 - G_i(z)) Q_i(z),
\end{aligned}$$

and evaluating at $\varepsilon = 0$:

$$\hat{R}''_z(\varepsilon) = (2g_i(z) + z g'_i(z)) Q_i(z) + \left(z g_i(z) - \frac{1}{2}(1 - G_i(z))\right) Q'_i(z),$$

and since $Q'_i(z) > 0$, this shows that

$$\hat{R}''_z(0) - R''_z(0) > 0.$$

Finally, observe that the countable set of mass points might be dense in W_i in which case there does not exist a small interval without a mass point and hence our initial hypothesis may never be satisfied. However, in this case, we can infer from Theorem 1 that the probability of $Q_i(z)$ to the right of every mass point w_i^k, must by equal to $Q_i\left(w_i^k\right)$, or $Q_i(z) = Q_i\left(w_i^k\right)$ for all $z \in (w_i^k, w_i^k + \varepsilon^k]$ and $\varepsilon^k > 0$ sufficiently small. But as the mass points are dense in W_i by hypothesis it follows that all points z with positive density belong to some interval of the form $(w_i^k, w_i^k + \varepsilon^k]$. Finally, notice that if the probability of receiving the object is identical for two realizations, then their virtual costs are identical, and hence it is optimal to bundle the realizations and represent them by their expected value. It thus follow that if the mass points were dense in W_i, it would be optimal to eliminate all realizations with positive density and bundle them with the (left-)adjacent mass point.■

The proof of Theorem 3 uses the following two auxiliary results.

Lemma 2 (Constant Winning Probabilities (I))
The optimal mechanism satisfies for all $\gamma_i^k, \gamma_i^{k+1}$ with $\gamma_i^k \geq \gamma_i^{k+1}$: $Q_i^k = Q_i^{k+1}$.

Proof. Suppose to the contrary (and by Theorem 1) that $Q_i^k < Q_i^{k+1}$. Then there must exist w_{-i} such that $q_i\left(w_i^k, w_{-i}\right) < q_i\left(w_i^{k+1}, w_{-i}\right)$. The incentive compatibility conditions of all agents except i, and in particular their expected winning probabilities remain constant under $q_i(\cdot)$ and a modified probability assignment $\widehat{q}_i(\cdot)$ as long as

$$f_i^k q_i\left(w_i^k, w_{-i}\right) + f_i^{k+1} q_i\left(w_i^{k+1}, w_{-i}\right) = f_i^k \widehat{q}_i\left(w_i^k, w_{-i}\right) + f_i^{k+1} \widehat{q}_i\left(w_i^{k+1}, w_{-i}\right). \tag{12}$$

Next observe that by the assumption of $\gamma_i^k \geq \gamma_i^{k+1}$, any $\hat{q}_i(\cdot)$ such that (12) is maintained and displays $q_i\left(w_i^k, w_{-i}\right) < \widehat{q}_i\left(w_i^k, w_{-i}\right)$ must (weakly) increase the revenues of the auctioneer, which delivers the contradiction. ■

In the next lemma we compare the revenues arising from two different partitions. To this end, we start with a given initial partition and then generate a second and modified partition by combining two adjacent partitional elements, k and $k+1$, of agent i into a single element.[4] The partitions are otherwise identical. The conditional expected value of the newly created element is given by:

$$\frac{w_i^k f_i^k + w_i^{k+1} f_i^{k+1}}{f_i^k + f_i^{k+1}}. \tag{13}$$

To better match up the original with the modified partition, it is convenient to keep the same number of partitional elements in the modified as in the original partition. We do this by simply assigning

[4] For the benefit of the reader, the same construction is also presented in the main body of the text, however there it appears in connection with Proposition 2 leading to Theorem 7.

the same conditional expected value (13) to element k and $k+1$, or

$$\bar{w}_i^k = \bar{w}_i^{k+1} \triangleq \frac{w_i^k f_i^k + w_i^{k+1} f_i^{k+1}}{f_i^k + f_i^{k+1}}, \tag{14}$$

where the upper bar always refers to values of the modified partition. The probability that the conditional expectation $\bar{w}_i^k = \bar{w}_i^{k+1}$ is realized in the modified partition is given by $f_i^k + f_i^{k+1}$. We then compare the revenue for the auctioneer under the original and the modified partition. As an aside, observe that we do not re-optimize the mechanism under the new information structure (by changing the probabilities of winning Q_i^{k-1}, Q_i^k and Q_i^{k+1}). The modified partition displays different virtual utilities only in the segments $k-1, k$, and $k+1$, where the local modification arises due to the combination of w_i^k and w_i^{k+1}, namely

$$\bar{\gamma}_i^{k-1} \triangleq w_{k-1}^i - \left(\bar{w}_i^k - w_i^{k-1}\right) \frac{1 - F_i^{k-1}}{f_i^{k-1}}, \tag{15}$$

and

$$\bar{\gamma}_i^k = \bar{\gamma}_i^{k+1} \triangleq \bar{w}_i^k - \left(w_i^{k+2} - \bar{w}_i^k\right) \frac{1 - F_i^{k+1}}{f_i^k + f_i^{k+1}}. \tag{16}$$

The difference in the revenues on any segment k is defined by

$$D_i^k \triangleq \left(\bar{\gamma}_i^k - \gamma_i^k\right) f_i^k. \tag{17}$$

Lemma 3 (Constant Winning Probabilities (II))
For all $w_i^{k-1} \leq w_i^k \leq w_i^{k+1} \leq w_i^{k+2}$ *and* $Q_i^{k-1} = Q_i^k = Q_i^{k+1} > 0:$

$$D_i^{k-1} Q_i^{k-1} + D_i^k Q_i^k + D_i^{k+1} Q_i^k = 0. \tag{18}$$

Proof. We first modify an arbitrary partition by joining two adjacent elements of the partition as suggested earlier in (13). The terms D_i^{k-1}, D_i^k and D_i^{k+1} are defined in (15)-(17), and can be written more explicitly after some cancellations as

$$D_i^{k-1} = \left(w_i^k - \bar{w}_i^k\right)\left(1 - F_i^{k-1}\right),$$

$$D_i^k = \left(\bar{w}_i^k - w_i^k\right)\left(1 - F_i^{k-1}\right) + \left(w_i^{k+2} - \bar{w}_i^k\right) \frac{\left(1 - F_i^{k-1}\right) f_i^{k+1}}{f_i^k + f_i^{k+1}} + \left(w_i^{k+1} - w_i^{k+2}\right)\left(1 - F_i^k\right),$$

and

$$D_i^{k+1} = \left(\bar{w}_i^k - w_i^{k+1}\right)\left(1 - F_i^k\right) + \left(w_i^{k+2} - \bar{w}_i^k\right) \frac{\left(1 - F_i^{k+1}\right) f_i^k}{f_i^k + f_i^{k+1}}.$$

The equality in (18) now follows directly after some elementary cancellations and using the decomposition of $\bar{w}_i^k$ given in (14). ■

Proof of Theorem 3. Suppose to the contrary and hence that there exists γ_i^k and γ_i^{k+1} such that $\gamma_i^k \geq \gamma_i^{k+1}$. By Theorem 1, Q_i^k has to be nondecreasing. By Lemma 2, the winning probabilities $q_i\left(\cdots, w_i^k, \cdots\right) = q_i\left(\cdots, w_i^{k+1}, \cdots\right)$ have to be identical for $\gamma_i^k \geq \gamma_i^{k+1}$. It thus follows that $Q_i^k = Q_i^{k+1}$. By Lemma 3, the revenue for the auctioneer remains unchanged when the mass point w_i^k and w_i^{k+1} are merged, provided that $Q_i^{k-1} = Q_i^k = Q_i^{k+1}$. Moreover as $D_i^{k-1} < 0$, it follows that the revenues are strictly improved if $Q_i^{k-1} < Q_i^k$. Thus it follows that every auction with non-monotone virtual utilities is (weakly) dominated by one with monotone virtual utilities.■

Proof of Corollary 1. The characterization follows immediately from pointwise optimization of the objective function (7) for any realization of values $w = \left(w_1^{k_1}, ..., w_I^{k_I}\right)$.■

Proof of Theorem 4. **(1.)** Suppose to the contrary. Then there exist γ_i^k such that

$$\left\{\gamma_j^k \left| \gamma_i^k < \gamma_j^k < \gamma_i^{k+1}\right.\right\} = \emptyset.$$

Observe next that if two adjacent virtual utilities belong to bidder i then the probability of receiving the good has to be identical on both intervals, $Q_i^k = Q_i^{k+1}$ by Theorem 3 and Corollary 1. But by the same argument as Theorem 3, we may then join the mass points w_i^k and w_i^{k+1} and the expected revenues for the auctioneer will strictly increase. A contradiction.

(2.) Suppose to the contrary and thus $\Gamma_i = \Gamma_j$ for all i, j. Then there exists *an* optimal auction such that for some i and some k, $Q_i^k = Q_i^{k+1}$. We can now appeal to the same argument as in (1.) to conclude that the revenues of the auctioneers are strictly increased by joining the mass points w_i^k and w_i^{k+1}, which destroys the symmetry in the virtual utilities.■

Proof of Theorem 5. First, we show that if the information structure is a partition, then the optimal information structure is a monotone partition. Consider a partitional point z_i separating S_i^k and S_i^l, where without loss of generality, $w_i^k < w_i^l$. A necessary condition for the optimality of z is

$$\frac{\partial R(\mathbf{z})}{\partial z_i} = 0.$$

The first order condition can be written more explicitly as

$$\begin{aligned} 0 \;=\; & \frac{\partial \gamma_i^{k-1}}{\partial z_i} f_i^{k-1} Q_i^{k-1} + \frac{\partial \gamma_i^k}{\partial z_i} f_i^k Q_i^k + \frac{\partial \gamma_i^{l-1}}{\partial z_i} f_i^{l-1} Q_i^{l-1} + \frac{\partial \gamma_i^l}{\partial z_i} f_i^l Q_i^l \\ & + \gamma_i^k f_i(z_i) Q_i^k - \gamma_i^l f_i(z_i) Q_i^l \\ & + \sum_{\left\{\gamma_j^k \left| \gamma_i^k < \gamma_j^k < \gamma_i^l \right.\right\}} \gamma_j^k f_j^k \prod_{m \neq i,j} \Pr\left(\gamma_m < \gamma_j^k\right) f_i(z_i). \end{aligned} \tag{19}$$

Using the composition of the virtual utility, γ_i^k, we can write (19) more explicitly as

$$\begin{aligned}
0 &= -\left(\frac{\left(z_i - w_i^k\right)\left(1 - F_i^{k-1}\right)}{f_i^k}\right) Q_i^{k-1} + \left(\left(\left(z_i - w_i^k\right) + \left(w_i^{k+1} - w_i^k\right)\right)\left(\frac{1 - F_i^{k-1}}{f_i^k}\right)\right) Q_i^k \\
&\quad + \left(w_i^{k+2} - w_i^{k+1}\right) Q_i^{k+1} + ... + \left(w_i^{l-1} - w_i^{l-2}\right) Q_i^{l-2} + \\
&\quad + \left(\frac{\left(z_i - w_i^l\right)\left(1 - F_i^{l-1}\right)}{f_i^l} + \left(w_i^l - w_i^{l-1}\right)\right) Q_i^{l-1} + \left(\left(-\left(z_i - w_i^l\right) + \left(w_{l+1}^i - w_i^l\right)\right)\left(\frac{1 - F_i^{l-1}}{f_i^l}\right)\right) Q_i^l \\
&\quad + \gamma_i^k Q_i^k - \gamma_i^l Q_i^l + \sum_{\left\{\gamma_j^k \middle| \gamma_i^k < \gamma_j^k < \gamma_i^l\right\}} \gamma_j^k f_j^k \prod_{m \neq i,j} \Pr\left(\gamma_m < \gamma_j^k\right).
\end{aligned}$$

We observe that we can rewrite the first order conditions as follows

$$a_i^{k,l} + \left(b_i^l - b_i^k\right) z_i = 0, \tag{20}$$

where

$$b_i^k = \frac{\left(Q_i^k - Q_i^{k-1}\right)\left(1 - F_i^{k-1}\right)}{f_i^k}, \tag{21}$$

and

$$b_i^l = \frac{\left(Q_i^l - Q_i^{l-1}\right)\left(1 - F_i^{l-1}\right)}{f_i^l}. \tag{22}$$

We now derive from (20)-(22) that the optimal partition must be monotone. First, we observe that $b_i^k \leq b_i^{k+1}$ for all k. By way of contradiction, suppose not. As $w_i^k < w_i^{k+1}$, there exist points $v_i \in S_i^k$ and $v_i' \in S_i^{k+1}$ such that $v_i < v_i'$. By the optimality conditions it has to be that

$$a_i^{k,k+1} + \left(b_i^{k+1} - b_i^k\right) v_i \leq 0 \text{ and } a_i^{k,k+1} + \left(b_i^{k+1} - b_i^k\right) v_i' \geq 0,$$

but this implies that $b_i^{k+1} \geq b_i^k$. Suppose for the moment then that $b_i^k < b_i^{k+1}$ holds as a strict inequality for all k. We can now view the (pairwise) first order conditions as a linear function of z_i as in (20), and it follows from the single crossing property of linear functions, that each element of the partition must consist of a single interval.

Finally, we observe that the argument above does not show that the partition must be monotone for $b_i^k = b_i^{k+1}$. However, if $b_i^k = b_i^{k+1}$, then a partition which is monotone must also be optimal (among possibly others). To see this, consider how the elements w_i^k and w_i^{k+1} appear in the revenue function of the auctioneer (see (10)):

$$w_i^k \left(1 - F_i^{k-1}\right)\left(Q_i^k - Q_i^{k-1}\right) + w_i^{k+1}\left(1 - F_i^k\right)\left(Q_i^{k+1} - Q_i^k\right),$$

or

$$\int_{S_i^k} v_i f(v_i)\, dv_i \frac{\left(1-F_i^{k-1}\right)}{f_i^k}\left(Q_i^k - Q_i^{k-1}\right) + \int_{S_{k+1}} v_i f(v_i)\, dv_i \frac{\left(1-F_i^{k}\right)}{f_i^{k+1}}\left(Q_i^{k+1} - Q_i^{k}\right). \tag{23}$$

But by assumption of $b_i^k = b_i^{k+1}$, it follows that

$$\frac{\left(1-F_i^{k-1}\right)}{f_i^k}\left(Q_i^k - Q_i^{k-1}\right) = \frac{\left(1-F_i^{k}\right)}{f_i^{k+1}}\left(Q_i^{k+1} - Q_i^{k}\right).$$

Thus we can infer from (23) that any distribution of points across S_i^k and S_i^{k+1} delivers the same value to the auctioneer provided that the density under S_i^k and S_i^{k+1} integrates up to f_i^k and f_i^{k+1}, respectively. But clearly one possible allocation is represented by a monotone partition such that $v_i \in S_i^k, v_i' \in S_i^{k+1}$ implies that $v_i < v_i'$.

Finally observe that the above argument immediately shows that the optimal information structure must be a partition rather than a noisy information structure. A noisy information structure would allocate the density of (at least) some realizations v_i across different signals s_i^k and associated expected valuations w_i^k. But from the single crossing properties of the first order conditions above, we can conclude that for all adjacent expected values, w_k^i and w_{k+1}^i, at most one realization v_i can optimally belong to either w_i^k and w_i^{k+1}. But this implies that with probability 1, a value v_i generates a single signal s_i^k and hence that the information structure is a partition.■

Proof of Theorem 6. Every monotone partition with at most K elements can be characterized as a $K+1$ dimensional vector, where each entry specifies a partitional point. The objective function of the auctioneer is continuous in the location of each partitional point. The space of possible locations for each partitional point is compact and hence by Weierstrass an optimal partition is guaranteed to exist.■

The proof of Proposition 1 uses the following auxiliary result for the behavior of the ratio

$$\frac{w_i^{k+1}\left(\frac{1-F_i^{k+1}}{F_i^{k+1}-F_i^k}\right) - z\left(\frac{1-F_i^{k}}{F_i^{k+1}-F_i^k}\right)}{\left(z-w_i^k\right)\left(\frac{1-F_i^{k-1}}{F_i^{k}-F_i^{k-1}}\right)},$$

which is central to the first order conditions for adjacent partitional elements. For a given partitional point z, consider adjacent elements of the partition described by $[z_i^{k-1}, z_i^k) \triangleq [z-\varepsilon^k, z)$ and $[z_i^k, z_i^{k+1}) \triangleq [z, z+\varepsilon^{k+1})$. We are interested in the behavior of the ratio as $\varepsilon^k, \varepsilon^{k+1} \to 0$.

Lemma 4 (Limit Ratio) *For any $z \in (0,1)$, the limit is given by:*

$$\lim_{\varepsilon^k, \varepsilon^{k+1} \to 0} \frac{w_i^{k+1}\left(\frac{1-F_i^{k+1}}{F_i^{k+1}-F_i^k}\right) - z\left(\frac{1-F_i^{k}}{F_i^{k+1}-F_i^k}\right)}{\left(z-w_i^k\right)\left(\frac{1-F_i^{k-1}}{F_i^{k}-F_i^{k-1}}\right)} = \frac{\frac{1}{2}\frac{1-F_i(z)}{f_i(z)} - z}{-\frac{1}{2}\frac{1-F_i(z)}{f_i(z)}}. \tag{24}$$

Proof. The first term in (24) can be written more explicitly as

$$\frac{\frac{\int_z^{z+\varepsilon^{k+1}} v_i f_i(v_i)dv_i}{\int_z^{z+\varepsilon^{k+1}} f_i(v_i)dv_i}\left(\frac{1-\int_0^{z+\varepsilon^{k+1}} f_i(v_i)dv_i}{\int_z^{z+\varepsilon^{k+1}} f_i(v_i)dv_i}\right)-z\left(\frac{1-\int_0^{z} f_i(v_i)dv_i}{\int_z^{z+\varepsilon^{k+1}} f_i(v_i)dv_i}\right)}{\left(z-\frac{\int_{z-\varepsilon^k}^{z} v_i f_i(v_i)dv_i}{\int_{z-\varepsilon^k}^{z} f_i(v_i)dv_i}\right)\left(\frac{1-\int_0^{z-\varepsilon^k} f_i(v_i)dv_i}{\int_{z-\varepsilon^k}^{z} f_i(v_i)dv_i}\right)}.$$

We first consider the limit as $\varepsilon^k, \varepsilon^{k+1} \to 0$. Consider first the numerator:

$$\lim_{\varepsilon^{k+1}\to 0}\left[\frac{\int_z^{z+\varepsilon^{k+1}} v_i f_i(v_i)\,dv_i}{\int_z^{z+\varepsilon^{k+1}} f_i(v_i)\,dv_i}\left(\frac{1-\int_0^{z+\varepsilon^{k+1}} f_i(v_i)\,dv_i}{\int_z^{z+\varepsilon^{k+1}} f_i(v_i)\,dv_i}\right)-z\left(\frac{1-\int_0^{z} f_i(v_i)\,dv_i}{\int_z^{z+\varepsilon^{k+1}} f_i(v_i)\,dv_i}\right)\right],$$

or equivalently

$$\lim_{\varepsilon^{k+1}\to 0}\left[\left(\frac{\int_z^{z+\varepsilon^{k+1}} v_i f_i(v_i)\,dv_i}{\int_z^{z+\varepsilon^{k+1}} f_i(v_i)\,dv_i}-z\right)\left(\frac{1-\int_0^{z} f_i(v_i)\,dv_i}{\int_z^{z+\varepsilon^{k+1}} f_i(v_i)\,dv_i}\right)-z\right].$$

As the left hand side of the first term converges to zero and the right hand side to infinity, we have to apply l'Hopital and evaluating the derivatives at $\varepsilon^{k+1}=0$ we get

$$\frac{1}{2}\frac{1-F_i(z)}{f_i(z)}.$$

Consider next the denominator, where we have

$$\lim_{\varepsilon^{k}\to 0}\left(z-\frac{\int_{z-\varepsilon^k}^{z} v_i f_i(v_i)\,dv_i}{\int_{z-\varepsilon^k}^{z} f_i(v_i)\,dv_i}\right)\left(\frac{1-\int_0^{z-\varepsilon^k} f_i(v_i)\,dv_i}{\int_{z-\varepsilon^k}^{z} f_i(v_i)\,dv_i}\right),$$

and similarly applying l'Hopital and evaluating the derivatives at $\varepsilon^k=0$ we get:

$$\frac{1}{2}\frac{1-F_i(z)}{f_i(z)}. \tag{25}$$

which concludes the proof. ∎

Proof of Proposition 1. As in Lemma 4, fix a partitional point z, and consider adjacent elements of the partition described by $[z_i^{k-1}, z_i^k) \triangleq [z-\varepsilon^k, z)$ and $[z_i^k, z_i^{k+1}) \triangleq [z, z+\varepsilon^{k+1})$. Consider the first-order conditions derived for Theorem 5 in equation (19), when specialized to the case of two adjacent partitional elements, or $l = k+1$:

$$\begin{aligned}0 \;=\;& \frac{\partial\gamma_i^{k-1}}{\partial z_i^k} f_i^{k-1}Q_i^{k-1}+\frac{\partial\gamma_i^{k}}{\partial z_i^k} f_i^{k}Q_i^{k}+\frac{\partial\gamma_i^{k+1}}{\partial z_i^k} f_i^{k+1}Q_i^{k+1}+\\ &+\gamma_i^k f_i(z)\,Q_i^k-\gamma_i^{k+1} f_i(z)\,Q_i^{k+1} \\ &+\sum_{\left\{\gamma_j^k \mid \gamma_i^k<\gamma_j^k<\gamma_i^{k+1}\right\}}\gamma_j^k f_j^k\prod_{m\neq i,j}\Pr\left(\gamma_m<\gamma_k^j\right)f_i(z).\end{aligned} \tag{26}$$

We observe initially that

$$\sum_{\left\{\gamma_j^k \mid \gamma_i^k<\gamma_j^k<\gamma_i^{k+1}\right\}} f_j^k\prod_{m\neq i,j}\Pr\left(\gamma_m<\gamma_j^k\right)=Q_i^{k+1}-Q_i^k.$$

We therefore have

$$\sum_{\left\{\gamma_j^k \middle| \gamma_i^k<\gamma_j^k<\gamma_i^{k+1}\right\}} \gamma_j^k f_j^k \prod_{m\neq i,j} \Pr\left(\gamma_m < \gamma_j^k\right) f_i\left(z\right) = \bar{\gamma}_i^k \left(Q_i^{k+1} - Q_i^k\right) f_i\left(z\right).$$

for some $\bar{\gamma}_i^k$ satisfying $\gamma_i^k < \bar{\gamma}_i^k < \gamma_i^{k+1}$. Consider next the remaining terms in (26). The partial derivatives can be computed explicitly and after combining terms, we can rewrite the first order conditions to collect the winning probabilities and get

$$\frac{Q_i^k - Q_i^{k-1}}{Q_i^{k+1} - Q_i^k} = \frac{w_i^{k+1}\frac{1-F_i^{k+1}}{f_i^{k+1}} - z\frac{1-F_i^k}{f_i^{k+1}} - \bar{\gamma}_i^k}{\left(z - w_i^k\right)\frac{1-F_i^{k-1}}{f_i^k}}.$$

The first-order condition has to hold for every (optimal) partition. Next we consider what happens to the term on the right hand side in the limit as $\varepsilon^k, \varepsilon^{k+1} \to 0$, or

$$\lim_{\varepsilon^k,\varepsilon^{k+1}\to 0} \frac{w_i^{k+1}\frac{1-F_i^{k+1}}{f_i^{k+1}} - z\frac{1-F_i^k}{f_i^{k+1}} - \bar{\gamma}_i^k}{\left(z - w_i^k\right)\frac{1-F_i^{k-1}}{f_i^k}}. \tag{27}$$

The term $\bar{\gamma}_k$ then converges to γ_i^k by a Sandwich argument and as

$$\lim_{\varepsilon^k,\varepsilon^{k+1}\to 0} \gamma_i^k = z - \frac{1 - F_i\left(z\right)}{f_i\left(z\right)},$$

we obtain from (27) that

$$\lim_{\varepsilon^k,\varepsilon^{k+1}\to 0} \frac{w_i^{k+1}\frac{1-F_i^{k+1}}{f_i^{k+1}} - z\frac{1-F_i^k}{f_i^{k+1}} - \bar{\gamma}_i^k}{\left(z - w_i^k\right)\frac{1-F_i^{k-1}}{f_i^k}} = 1.$$

The ratio is continuous in z for every given i, ε^k, and ε^{k+1} as the density $f_i\left(z\right)$ is continuous in z by assumption. As z is element of a compact set and there are only a finite number of bidders, it follows that the limit result can be extended uniformly for all z and i, provided that the intervals $[z_i^{k-1}, z_i^k) \triangleq [z - \varepsilon^k, z)$ and $[z_i^k, z_i^{k+1}) \triangleq [z, z + \varepsilon^{k+1})$ are sufficiently small or $\varepsilon_i^k, \varepsilon_i^{k+1} \leq \varepsilon$ for all i and k.■

Proof of Proposition 2. This argument is a continuation of Lemma 3. As $D_i^{k-1} < 0$, it follows that if $D_i^{k+1} \geq 0$, then

$$D_i^{k-1}Q_i^{k-1} + D_i^k Q_i^k + D_i^{k+1}Q_i^{k+1} > 0, \tag{28}$$

as

$$Q_i^{k-1} < Q_i^k < Q_i^{k+1},$$

from the optimality of the auction. Next we want to show that the inequality also holds for $D_i^{k+1} < 0$. As in Proposition 1, fix a partitional point z, and consider adjacent elements of the partition described by $[z_i^{k-1}, z_i^k) \triangleq [z - \varepsilon^k, z)$ and $[z_i^k, z_i^{k+1}) \triangleq [z, z + \varepsilon^{k+1})$ and let $\varepsilon^k, \varepsilon^{k+1} \to 0$. By Proposition 1,

$$\frac{Q_i^k - Q_i^{k-1}}{Q_i^{k+1} - Q_i^k} \to 1.$$

It is therefore sufficient to show that

$$\lim_{\varepsilon^k, \varepsilon^{k+1} \to 0} \frac{D_i^{k+1}}{D_i^{k-1}} < 1 - \delta,$$

holds for some $\delta > 0$. In fact, we next show that

$$\lim_{\varepsilon^k, \varepsilon^{k+1} \to 0} \frac{D_i^{k+1}}{D_i^{k-1}} \leq 0. \tag{29}$$

As the limit of the ratio may depend on the rate at which ε^k and ε^{k+1} converge to zero, we have to show that the inequality (29) holds for any sequence of ε^k and ε^{k+1}. To this end, define an arbitrary smooth path,

$$h : [0, 1] \to \mathbb{R}_+^2,$$

with

$$h : t \longmapsto \left(\varepsilon^k(t), \varepsilon^{k+1}(t)\right),$$

such that $\varepsilon^k(0) = \varepsilon^{k+1}(0) = 0$, $\varepsilon^{k\prime}(0)$, $\varepsilon^{k+1\prime}(0) > 0$ and $\varepsilon^k(t) > 0$, $\varepsilon^{k+1}(t) > 0$ for all $t > 0$. Define also

$$\hat{D}_i^{k+1} \triangleq \frac{\left(z_i^{k+1} - \bar{w}_i^k\right) \frac{\left(1 - F_i^{k+1}\right) f_i^k}{f_i^k + f_i^{k+1}} - \left(w_i^{k+1} - \bar{w}_i^k\right)\left(1 - F_i^k\right)}{\left(w_i^k - \bar{w}_i^k\right)\left(1 - F_i^{k-1}\right)},$$

where $\hat{D}_i^{k+1}$ differs from D_i^{k+1} only insofar as we replaced w_i^{k+2} by z_i^{k+1}. It therefore follows that

$$\frac{\hat{D}_i^{k+1}}{D_i^{k-1}} < \frac{D_i^{k+1}}{D_i^{k-1}}.$$

For a fixed path $h(\cdot)$, consider then the limit

$$\lim_{t \to 0} \frac{\hat{D}_i^{k+1}(t)}{D_i^{k-1}(t)}.$$

As both terms in the ratio converge to zero, we have to apply l'Hopital, or

$$\lim_{t \to 0} \frac{\hat{D}_i^{k+1}(t)}{D_i^{k-1}(t)} = \lim_{t \to 0} \frac{\frac{\partial \hat{D}_i^{k+1}}{\partial \varepsilon^k} \frac{d\varepsilon^k}{dt} + \frac{\partial \hat{D}_i^{k+1}}{\partial \varepsilon^{k+1}} \frac{d\varepsilon^{k+1}}{dt}}{\frac{\partial D_i^{k-1}}{\partial \varepsilon^k} \frac{d\varepsilon^k}{dt} + \frac{\partial D_i^{k-1}}{\partial \varepsilon^{k+1}} \frac{d\varepsilon^{k+1}}{dt}}.$$

Evaluating the partial derivatives at $\left(\varepsilon^k, \varepsilon^{k+1}\right) = (0, 0)$, and using the limit properties of the conditional means as recorded in Lemma 1, results in

$$\lim_{t \to 0} \frac{\frac{1}{2}\frac{d\varepsilon^k}{dt} - \left(\frac{1}{2}\frac{d\varepsilon^k}{dt} - 0\frac{d\varepsilon^{k+1}}{dt}\right)}{\left(0\frac{d\varepsilon^k}{dt} - \frac{1}{2}\frac{d\varepsilon^{k+1}}{dt}\right)} = -\frac{0}{\frac{1}{2}\frac{d\varepsilon^{k+1}}{dt}}.$$

As

$$\frac{d\varepsilon^{k+1}}{dt} > 0,$$

at $t = 0$, it follows that the limit is unique and equal to zero, which completes the proof.■

Proof of Theorem 7. Consider first an optimal information structure for the auctioneer subject to the restriction that the information structure forms a monotone partition with up to K elements in every partition. Denote the resulting revenue for the auctioneer by $R(K)$. Without loss of generality, we may assume that the partition of each agent i has at most one element, namely $k_i = 1$, which has a zero probability of winning ($Q_i^1 = 0$). We can make this assumption by the earlier results of Theorem 3 and 5. The argument is now by contradiction. Suppose there doesn't exists an optimal finite partition. Then for every K, there must exist a K' such that

$$R(K') > R(K) \text{ with } K < K',$$

and in fact

$$R(K') > R(K'') \text{ for all } K'' < K'. \tag{30}$$

Next we observe that by Lemma 2, there exists an $\bar{\varepsilon}$ such that a coarser information structure (by combining adjacent intervals smaller than $\bar{\varepsilon}$) yields a higher revenue. But for every $\bar{\varepsilon} > 0$, there exists $K(\bar{\varepsilon})$ such that if the partition of agent i has $K(\bar{\varepsilon}) < \infty$ or more elements, there must at least be one pair of adjacent elements, where the size of each interval is less than $\bar{\varepsilon}$. But now suppose that $K' > K(\bar{\varepsilon})$, then Lemma 2 tells us that

$$R(K') < R(K' - 1),$$

which delivers the desired contradiction. ■

Proof of Theorem 8. We argue by induction. First, we show that, as $I \to \infty$, there exists a subsequence $I_K \to \infty$, such that all $i \in I_K$ have an interval of length ε at the top. Second, we show that if there exists a subsequence, $I_k \to \infty$, such that all $i \in I_k$ have an interval at $[(k-1) \cdot \varepsilon, k \cdot \varepsilon)$, then there exists a subsequence $I_{k-1} \to \infty$ of the sequence I_k such that all $i \in I_{k-1}$ have an interval at $[(k-2) \cdot \varepsilon, (k-1) \cdot \varepsilon)$. The induction argument implies that the exists subsequence $I_1 \to \infty$ with the property that all $i \in I_1$ have an ε-partition.

First, suppose to the contrary. That is there exists a finite M such that at most M bidders have an interval $[1-\varepsilon, 1]$ in their partition as $I \to \infty$. This implies that at most M bidders have a valuation of $1-\frac{\varepsilon}{2}$. Moreover, at least one of these M bidders will have a valuation of $1-\frac{\varepsilon}{2}$ with probability $1-F(1-\varepsilon)^M$. The remaining bidders have a highest valuation of at most $1-\varepsilon$. Thus, with probability $F(1-\varepsilon)^M$ the highest possible revenue equals $1-\varepsilon$. We may write an upper bound for the revenues as,

$$R_M \leq (1-\frac{\varepsilon}{2})[1-F(1-\varepsilon)^M] + (1-\varepsilon)F(1-\varepsilon)^M,$$

where the bound equals $(1-\frac{\varepsilon}{2})$ if at least one of M bidders has a valuation of $(1-\frac{\varepsilon}{2})$ which occurs with probability $1-F(1-\varepsilon)^M$. Otherwise, if none of the M bidders has a valuation of $(1-\frac{\varepsilon}{2})$, then all the remaining (virtual) valuations equal at most $(1-\varepsilon)$ and this event occurs with probability $F(1-\varepsilon)^M$.

Now, consider the revenues under the mechanism in which bidders receive the good only if their valuation falls into the smallest possible interval at the top. Clearly, if I_K bidders have a partition with the smallest possible interval at the top, then with probability $1-F(1-\varepsilon)^{I_K}$ the revenue equals $1-\frac{\varepsilon}{2}$. The revenues under this mechanism are at least,

$$R \geq (1-\frac{\varepsilon}{2})[1-F(1-\varepsilon)^{I_K}].$$

Combining the above inequalities yields,

$$R - R_M \geq \frac{\varepsilon}{2}F(1-\varepsilon)^M - (1-\frac{\varepsilon}{2})[F(1-\varepsilon)^{I_K}].$$

For I_K sufficiently large the right hand side is positive. A contradiction. Thus it cannot be that there exists a finite M such that at most M bidders have an interval $[1-\varepsilon, 1]$ in their partition as $I \to \infty$. In other words, it has to be that I_K bidders have an interval of length ε at the top with $I_K \to \infty$, as $I \to \infty$.

Second, suppose there exists a subsequence, $I_k \to \infty$, such that all $i \in I_k$ have an interval at $[(k-1)\cdot\varepsilon, k\cdot\varepsilon)$. We argue that this implies that there exists a subsequence $I_{k-1} \to \infty$ of the sequence I_k such that all $i \in I_{k-1}$ have an interval at $[(k-2)\cdot\varepsilon, (k-1)\cdot\varepsilon)$. The virtual valuation of agents $i \in I_k$ equals $w_k^i + (w_{k+1}^i - w_k^i)\frac{1-F_i^k}{f_i^k}$. Observe that the virtual utility depends on the length of the interval. Clearly, the virtual valuation is maximized if the interval is small which arises if the interval $[(k-1)\cdot\varepsilon, k\cdot\varepsilon)$ is contained in the partition of agent i. Let γ^k denote this virtual valuation. We can write the seller's revenue as a sum of two parts: First, the revenues conditional on virtual valuations exceeding γ^k or being equal to γ^k for $i \notin I_k$ and, second, the revenues conditional on virtual valuations being at most as large as γ^k,

$$R = \sum_{\{\gamma_i^k|\gamma_i^k>\gamma^k,\ \forall i\in I,\ \text{or}\ \gamma_i^k=\gamma^k\ \text{for}\ i\notin I_k\}} \left[\max(\gamma_i^k,0) f_i^k \prod_{j\neq i} F_j^{k(i)}\right] + \sum_{\{\gamma_i^k|\gamma_i^k\leq\gamma^k,\ \forall i\in I,\ \text{or}\ \gamma_i^k=\gamma^k\ \text{for}\ i\notin I_k\}} \left[\max(\gamma_i^k,0) f_i^k \prod_{j\neq i} F_j^{k(i)}\right]$$

where $F_j^{k(i)}$ is defined as in (9).

To maximize the seller's revenues it has to be that the second part in the revenues, which we denote by R_2, is maximized. In other words, revenues conditional on the event that all virtual valuations are below γ^k have to be maximized. Clearly if $\gamma^k \leq 0$, then the revenues are not affected by the choice of intervals and we are done. So, suppose that $\gamma^k > 0$. We show that to maximize R_2 there has to be a subsequence $I_{k-1} \to \infty$ of the sequence I_k such that all $i \in I_{k-1}$ have an interval at $[(k-2)\cdot\varepsilon,(k-1)\cdot\varepsilon)$. Suppose to the contrary, that is there exists a finite M such that at most M bidders have an interval $[(k-2)\cdot\varepsilon,(k-1)\cdot\varepsilon)$ in their partition as $I_k \to \infty$. This implies that at most M bidders have a virtual valuation of γ_k. Moreover, at least one of these M bidders will have a valuation of γ_k with probability $[[F(k\cdot\varepsilon)-F((k-1)\cdot\varepsilon)]/F(k\cdot\varepsilon)]^M$. Due to the discreteness of virtual valuations, there exists a $\delta > 0$ such that all remaining bidders have a highest valuation of at most $\gamma_k - \delta$. Thus, with probability $[[F(k\cdot\varepsilon)-F((k-1)\cdot\varepsilon)]/F(k\cdot\varepsilon)]^M$ the highest possible revenue equals $\gamma_k - \delta$. We may write an upper bound for the revenues as,

$$R_M \leq \gamma_k[1-\{\frac{F((k-1)\cdot\varepsilon)}{F(k\cdot\varepsilon)}\}^M] + (\gamma_k-\delta)\{\frac{F((k-1)\cdot\varepsilon)}{F(k\cdot\varepsilon)}\}^M$$

Now, consider the revenues under the mechanism in which bidders receive the good only if their valuation falls into the interval $[(k-1)\cdot\varepsilon, k\cdot\varepsilon)$. Clearly, if I_k bidders have a partition with an interval $[(k-1)\cdot\varepsilon, k\cdot\varepsilon]$, then with probability $[1-\{F((k-1)\cdot\varepsilon)/F(k\cdot\varepsilon)\}^{I_{k-1}}]$ the revenue equals γ_k. The revenues under this mechanism are at least,

$$R \geq \gamma_k[1-\{\frac{F((k+1)\cdot\varepsilon)}{F(k\cdot\varepsilon)}\}^{I_{k-1}}]$$

Combining the above inequalities yields,

$$R - R_M \geq \delta\cdot\{\frac{F((k-1)\cdot\varepsilon)}{F(k\cdot\varepsilon)}\}^M - \gamma_k[\{\frac{F((k-1)\cdot\varepsilon)}{F(k\cdot\varepsilon)}\}^{I_{k-1}}]$$

For I_{k-1} sufficiently large the right hand side is positive. A contradiction. Thus, it cannot be that there exists a finite M such that at most M bidders have an interval $[(k-2)\cdot\varepsilon,(k-1)\cdot\varepsilon)$ in their partition as $I_k \to \infty$. In other words, it has to be that I_{k-1} bidders have such an interval with $I_{k-1} \to \infty$, as $I_k \to \infty$. By induction we have shown that there exists a sequence I_1 such that all bidders $i \in I_1$ have an ε-partition.$\blacksquare$

Proof of Theorem 9. By Theorem 7, the set of virtual utilities is finite for every bidder i. By Theorem 3, the virtual utilities are strictly increasing for every bidder i. Thus the auctioneer can implement the optimal auction by offering the object according to the order of virtual utilities as represented in (8). The exclusive offer to agent i with realization w_i^k would carry the price T_i^k/Q_i^k. Finally, by Theorem 4, no bidder has two adjacent utilities in (8) and hence no bidder will receive two subsequent offers.$\blacksquare$

References

Bergemann, D. & J. Valimaki. forthcoming. "Information Acquisition and Efficient Mechanism Design." *Econometrica* .

Crawford, V. & J. Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50:1431–1452.

Cremer, J. & F. Khalil. 1992. "Gathering Information Before Signing a Contract." *American Economic Review* 82:566–578.

Cremer, J., F. Khalil & J-C. Rochet. 1998*a*. "Contracts and Productive Information Gathering." *Games and Economic Behavior* 25:174–193.

Cremer, J., F. Khalil & J-C. Rochet. 1998*b*. "Strategic Information Gathering Before a Contract is Offered." *Journal of Economic Theory* 81:163–200.

Laffont, J.J. 1989. *The economics of uncertainty and information.* Cambridge: MIT Press.

Lewis, T. R. & D. E. M. Sappington. 1993. "Ignorance in Agency Problems." *Journal of Economic Theory* 61:169–183.

Mas-Collel, A., M.D. Whinston & J.R. Green. 1995. *Microeconomic Theory.* Oxford: Oxford University Press.

Matthews, S. 1977. "Information Acqusition in Competitive Bidding Process." California Institute of Technology.

Matthews, S. 1984. Information Acquisition in Discriminatory Auctions. In *Bayesian Models in Economic Theory*, ed. M. Boyer & R.E. Kihlstrom. Amsterdam: North-Holland pp. 181–207.

Milgrom, P.R. & R.J. Weber. 1982. "A Theory of Auctions and Competitive Bidding." *Econometrica* 50:1089–1122.

Myerson, R. 1981. "Optimal Auction Design." *Mathematics of Operations Research* 6:58–73.

Perry, M. & P. J. Reny. 1999. "On the Failure of the Linkage Principle in Multi-Unit Auctions." *Econometrica* 67:895–900.

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

Sobel, J. 1993. "Information Control in the Principal-Agent Problem." *International Economic Review* 34:259–269.

Stegeman, M. 1996. "Participation Costs and Efficient Auctions." *Journal of Economic Theory* 71:228–259.

Tan, G. 1992. "Entry and R and D in Procurement Contracting." *Journal of Economic Theory* 58:41–60.

Figure 1: Optimal Partitions

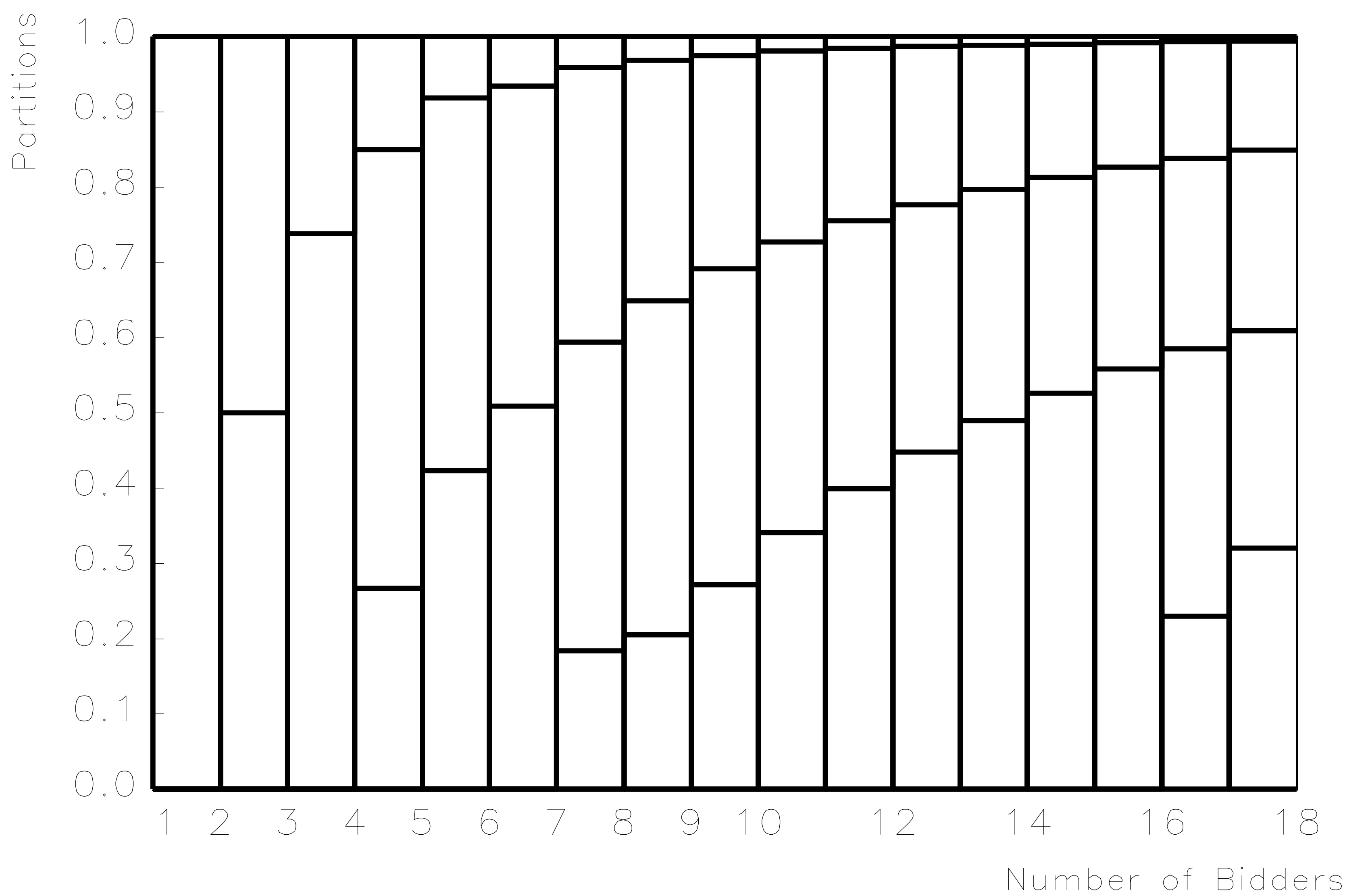

Figure 2: Optimal Virtual Valuations

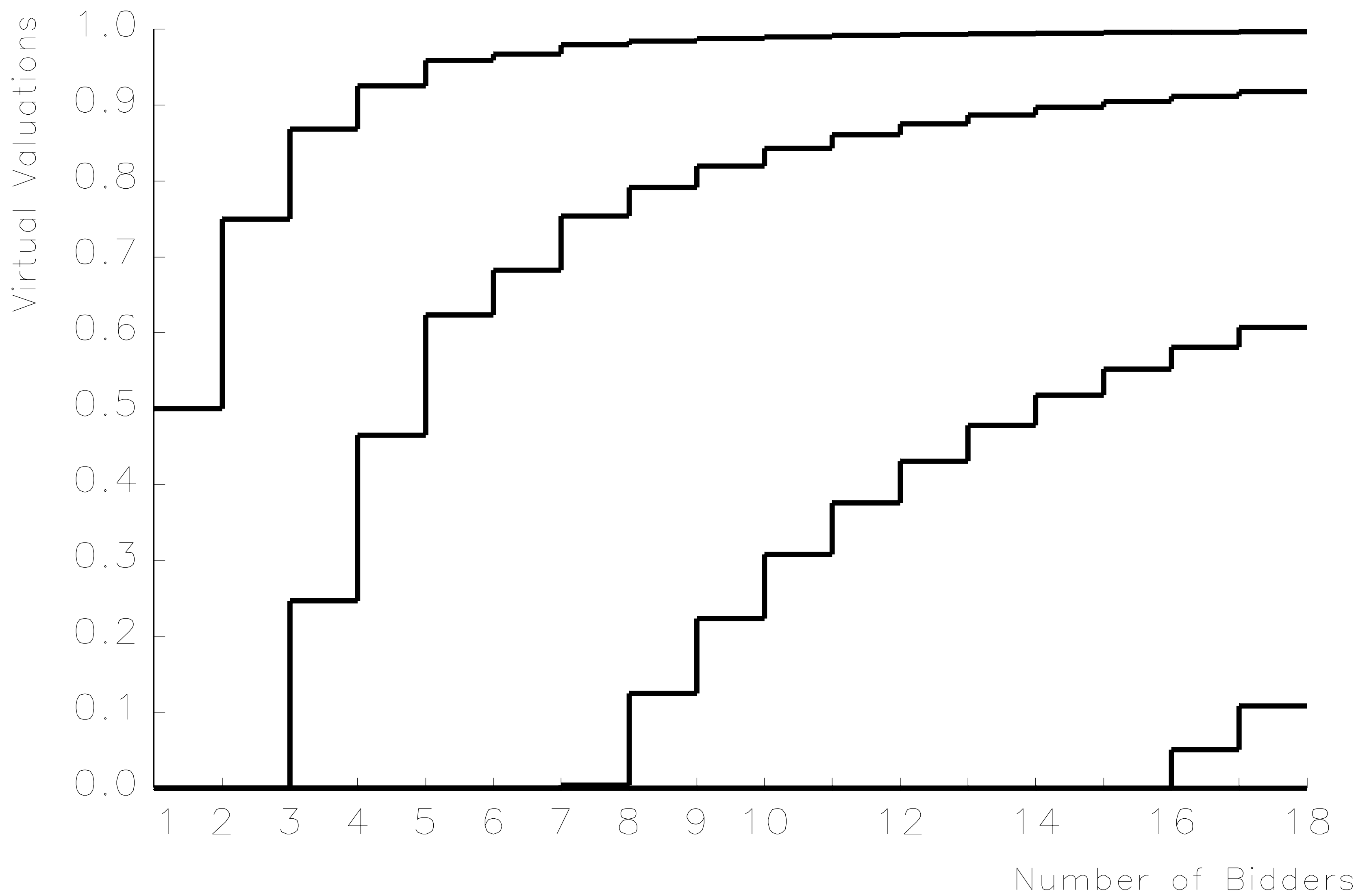

Figure 3: Revenues

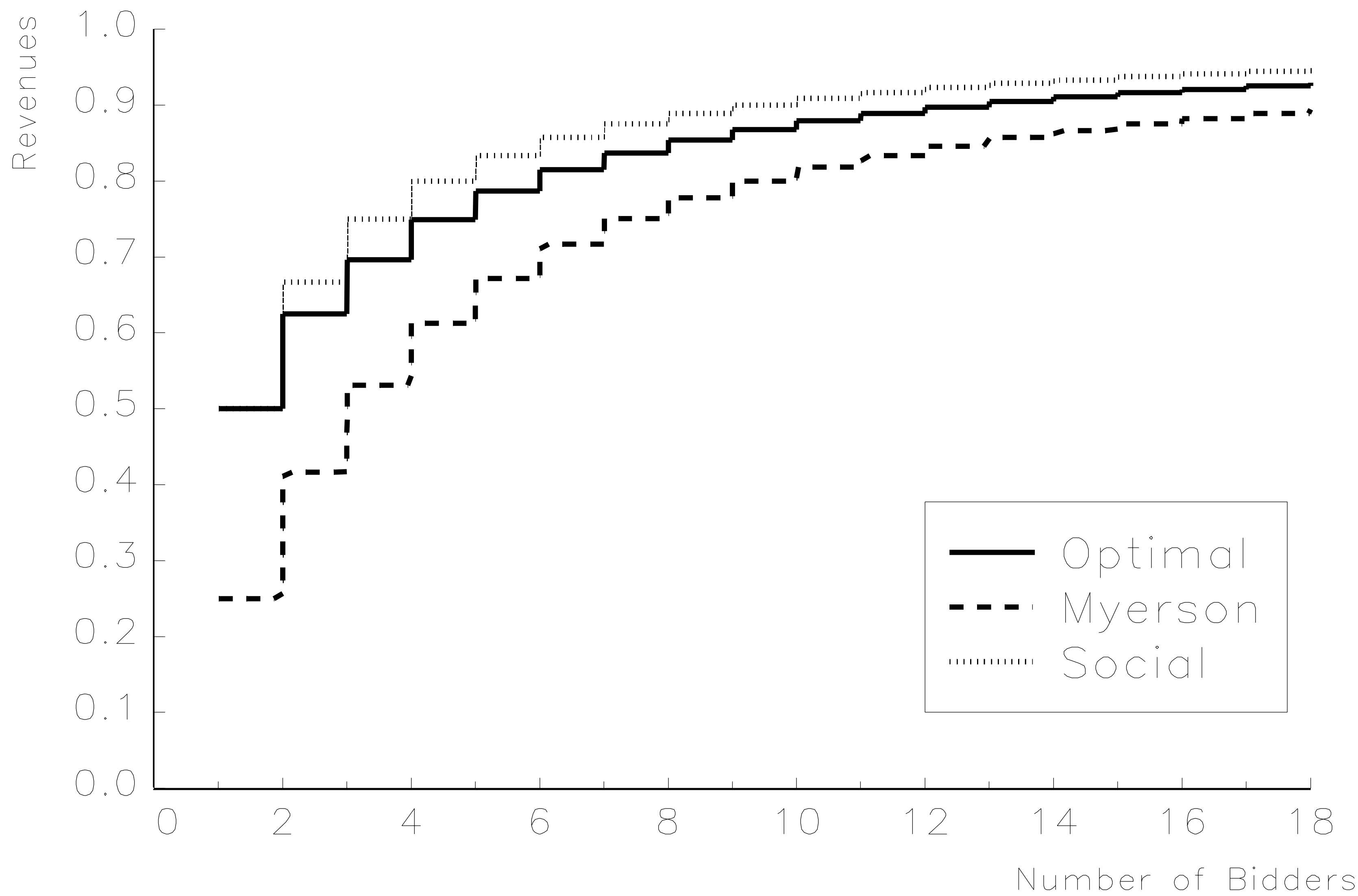

Figure 4: Relative Revenue Loss