Information Aggregation and the information content of order statistics

By Ilan Kremer and Andrzej Skrzypacz

This version
September, 16, 2003

ABSTRACT

We examine the link between the statistical properties of order statistics and information aggregation in auctions. We show how this link can be used in obtaining results in two opposite directions: (i) one can use auction theory in the analysis of order statistics, and, (ii) one can use results about order statistics to learn about information aggregation in auction.

In particular, we show how one can use results from auction theory in ranking the limiting information content of order statistics. We use a thought experiment in which we allocate the signals to bidders who compete in a $k + 1$-price auction. We show using economic arguments that the amount of information in the k-th order statistics is increasing in k in the limit when the number of signals increases and k is fixed. Still, even the first-order statistic contains non-trivial information.

Next, we use the properties of order statistics to show that information aggregation may not hold in a general model of affiliation. That is, information aggregation may critically depend on the assumption of conditional independence, which is the basis to what is known as the “Mineral Rights” model.

Introduction

Recently there has been a growing interest in the properties of common value auctions with a large number of bidders. This continues a classic line of research that examines markets with many strategic agents. Such analysis provides insights to the way prices aggregate private information and hence gives foundation to the concept of Rational Expectation Equilibrium.

The properties of prices in these auctions are strongly related to the statistical properties of order statistics. The price in a $k+1$-price auction is a function of the $k+1$ order statistic and hence does not contain more information. In fact, in many cases the price simply converges to the expected value of the asset conditional on this order statistic. In this paper we show how one can use this link in two ways. One can obtain results about information aggregation using results about order statistics and vice versa.

First, we consider a collection of signals $\{S_i\}_{i=1}^{n}$ that are i.i.d. conditional on some state variable, V where all random variables are distributed on $[0,1]$ according to an atomless distribution. We are interested in determining what the information content is of the k-th order statistic for a fixed k when we assume that n is large. To answer this question we conduct the following thought experiment: we assign these signals to agents and run a $k+1$-price auction in which bidders compete for k identical goods whose value is given by V. When the number of signals/agents is large, agents compete away their surplus. Using this and the form of the the bidding strategy in a $k+1$-price auction, we show that the information content of order statistics can be ranked.

One may guess that since fixed order statistics are likely to converge to one (the highest possible signal) independent of V, they do not convey any information. This is likely to be the case with discrete signals, e.g. binary signals. When signals have an atom less distribution, this does not take into account that convergence rate may depend on V. As a result these random variables convey non-trivial information even in the limit. In particular, we show that: (i) For a fixed k, as n grows to infinity, the $k+1$ order statistic contains in the limit strictly more information than the k-th order statistic. That is, the second-order statistic is more informative than the first-order statistic while the third-order statistic is more informative than the second-order statistic, etc. (ii) While the first-order statistic is less informative than any other order statistic, it contains non-trivial information even in the limit.

Next we demonstrate that information aggregation that was demonstrated in Pesendorfer and Swinkels (1997) (PS hereafter) may critically depend on the conditional independence assumption. The general information framework used in auction theory is that of affiliated random variables (see Milgrom and Weber (1982)). Affiliated random variables can be roughly described as random variables that are positively correlated even when conditioning on specific events. A special case is the ‘Mineral Rights’ model, which is commonly used in the literature on information aggregation. This model does not allow for signals to be correlated once we condition on the asset’s value. Such correlation may be present when agents are subject to a common source of noise or if agents rely on the same source of information. As we demonstrate, such a setup may satisfy affiliation and the collection of signals enables exact estimation of the asset’s value. However, in contrast to PS, a k_n+1-price auction with supply proportional to the number of agents fails to aggregate information. Hence, the conditional independence may not be an innocuous assumption. This result demonstrates the weakness of static auctions in aggregating information. A dynamic game such as the English auction aggregates information even in this case as prices depend on almost all signals and not just a single order statistic.

The information in order statistics

Consider a random variable, $V,$ distributed on $[0,1]$ according to a $g(v).$ We try to estimate V

using signals that are given by $\{S_i\}_{i=1}^{\infty}$. We assume that $\{S_i\}_{i=1}^{\infty}$ are also distributed on $[0,1]$; conditional on V, signals are distributed i.i.d according to $f(s|v)$. We assume that $\{S_i\}_{i=1}^{\infty}$ and V are affiliated so that a higher signal S_i implies a higher conditional expected value for V. We examine what can be inferred about V when observing the k-th order statistic. We define $Y_n(k)$ to be the $k-th$ order statistic among the first n signals, $\{S_i\}_{i=1}^{n}$. Our main result is that using auction theory we can rank the limiting information in order statistics. We illustrate the results with the following example:

Example *Consider the case where $V = \{0,1\}$ and*

$$f(s|v) = \left\{ \begin{array}{c} 1 \text{ if } v = 0 \\ 2s \text{ if } v = 1 \end{array} \right\}$$

Suppose there are $n = 1000$ signals and we observe the highest or the fifth highest signal. To what extent can we tell what the value of V is? Using Bayes rule we can express the expected value of V conditional on observing the $l-th$ highest signal, $E(V|Y_n(l) = s))$:

$$\frac{\int vF(S_i = s|v)^{n-l} f(S_i = s|v)(1 - F(S_i = s|v))^{l-1} f(v)dv}{\int F(S_i = s|v)^{n-l} f(S_i = s|v)(1 - F(S_i = s|v))^{l-1} f(v)dv}$$

In Figure 1 we use a Monte-Carlo simulation to plot the density function of random variable $E(V|Y_n(1))$, that is the expected value of the asset conditional on the first-order statistic. Figure 2 shows the results for the fifth-order statistic, that is, $E(V|Y_n(5))$. In both cases we plot two curves that represent the conditional densities of the random variables when we condition on V being 0 and 1.

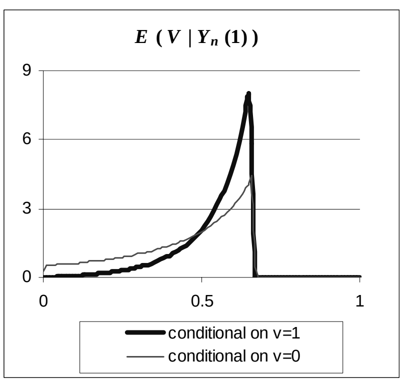

Figure 1

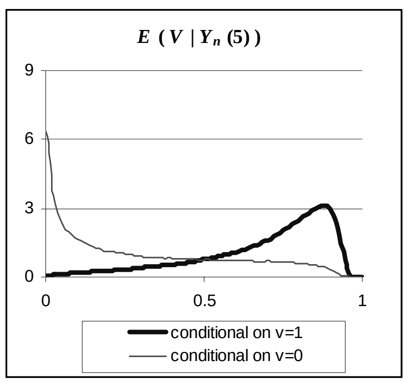

Figure 2

Two conclusions can be made:

- *The fifth-order statistic is more informative than the first-order statistic. The two conditional distributions are more seperated when we use the fifth order statistic.*
- *Even the first-order statistic contains some information. This can be concluded from the fact that the cumulativ distribution function $E(V|Y_n(1))$ is not a step function around 0.5. Moreover, the two conditional distributions are not identical.*

For the general result we make the following assumptions:

(A1) For any $\varepsilon > 0$ there exists $\delta > 0$ such that if $|s - s'| < \delta$ then for almost every v

$$\left| \frac{f(v|s)}{f(v|s')} - 1 \right| < \varepsilon.$$

(A2) Monotone likelihood ratio (MLRP) $\frac{f(s|v)}{f(s'|v)} > \frac{f(s|v')}{f(s'|v')}$ for $s > s', v > v'$.

(A3) There exists some $a, b > 0$ so that for any s, v and v' we have that: $a < f(s|v) < b$ and $\frac{f(v)}{f(v')} > a$.

(A1) is a uniform continuity condition in signals that applies across v. It implies that nearby signals provide similar information about the realization of V. (A2) means that s and v are strictly affiliated. This is a standard assumption in the auction literature and it implies a positive correlation between signals and the value of the asset. Lastly, assumption (A3) implies that there is a limited amount of information in every signal. While in the limit the collection of all signals is sufficient to determine V_n, a single signal is never sufficient. It implies a finite likelihood ratio of V given S_i. See PS for a more detailed discussion of the implications of such bounds.

Definition *(i) We say that a sequence of random variables $\{W_i\}$ is at least as informative as $\{Z_i\}$ in the limit if $E(X|W_n, Z_n) - E(X|W_n) \to 0$ in probability*
(ii) We say that a sequence of random variables $\{W_i\}$ is more informative than $\{Z_i\}$ in the limit if it is at least as informative as $\{Z_i\}$, and, $E(X|Z_n) - E(X|W_n) \nrightarrow 0$ in probability.
(iii) We say that a sequence of random variables $\{W_i\}$ contains non-trivial information in the limit if $E(X) - E(X|W_n) \nrightarrow 0$

We first argue that:

Lemma *Assuming A1-3, then $E(V|Y_n(k+1)) - E(V|Y_n(k)) \nrightarrow 0$ in probability.*

We provide the formal proof which is a bit technical in the appendix. A short outline is as follows. We first note that since signals are in $[0,1]$, the maximal conditional expectations for the value of the asset are given by: $E(V|Y_n(k+1) = 1)$, $E(V|Y_n(k) = 1)$. It also follows that both $E(V|Y_n(k+1) = 1)$ and $E(V|Y_n(k) = 1)$ do not depend on n and $E(V|Y_n(k+1) = 1) > E(V|Y_n(k) = 1)$. The claim then follows from the fact that we can find some threshold s_n^* close to one so that (i) $E(V|Y_n(k+1) = s) > E(V|Y_n(k) = 1)$ for any $s > s_n^*$, and (ii) $\Pr(|Y_n(k+1) > s_n^*)$ is bounded away from zero.

Theorem *Assume A1-3. For any k the $k+1$ order statistic, $Y_n(k+1)$, is more informative in the limit than the k order statistic, $Y_n(k)$.*

Proof Consider the $k+1$-price auction and let P_n denote the resulting price. That is there are k units to be sold and the price is the $k+1$-highest bid. The bidding strategy for an agent with a signal s in this case is given by:

$$b_n(s) = E(V|Y_n(k) = Y_n(k+1) = s)$$

The price is given by the random variable $P_n = b_n(Y_n(k+1))$. Note that

- $b_n(Y_n(k+1)) \leq E(V|Y_n(k+1))$ almost surely.
- $b_n(Y_n(k+1)) \leq E(V|Y_n(k+1), Y_n(k))$ almost surely.

Jackson and Kremer (2002) prove that in this case the auction is competitive so that $E(V) - E(P_n) \to 0$. This implies:

$$P_n - E(V|Y_n(k+1)) \to 0 \text{ in probability}$$

$$P_n - E(V|Y_n(k+1), Y_n(k)) \to 0 \text{ in probability}$$

Therefore we conclude that $E(V|Y_n(k+1)) - E(V|Y_n(k+1), Y_n(k)) \to 0$ in probability. This implies that $Y_n(k+1)$ is at least as informative as $Y_n(k)$. The Theorem then follows from Lemma ref: diff.

Theorem *The first-order statistic $Y_n(1)$ contains non-trivial information.*

Proof Consider the first-price auction and let P_n denote the resulting price. In this auction the price satisfies $P_n \leq E(V|Y_n(1))$ almost surely. Since $E(P_n) - E(V) \to 0$ we conclude that $P_n - E(V|Y_n(1)) \to 0$ in probability. Now assume by contradiction that $P_n - E(V) \to 0$. In this case with positive probability there is an agent with a high signal such that $E(V|S_i) > E(V) + \varepsilon$. This agent can bid $E(V) + \varepsilon/2$ and ensure a profit that is bounded away from zero.

Affiliation, Mineral Rights and Information Aggregation

Consider a sequence of auctions with an increasing number of bidders in which supply is proportional to the number of bidders. In particular, we examine a $k_n + 1$ – price auction, in which the supply is proportional to the number of bidders, $k_n/n \to \alpha \in (0,1)$. PS have shown that if signals are conditionally independent (conditional on the asset's value) and identically distributed then the price aggregates information. This setup of conditional independence is known as the 'Mineral Rights' model and is commonly used in the literature. As we discussed in the introduction, this is a special case of affiliation that was introduced in Milgrom and Weber (1982). It does not allow for correlation among signals once we condition on the value of the asset. If instead agents are also subject to some common noise, their signals in general could still be affiliated even though they would not satisfy the Mineral Rights assumption. Such common noise may be a result of common mistakes, reliance on a common source of information, error in measurements etc. As we shall demonstrate, while the aggregate information may be sufficient for exact estimation of the asset's value, the $k_n + 1$ – price auction does not aggregate information. Thus, the assumption of conditional independence is not an innocuous one.

To see this explicitly, suppose that the value of the asset, V, is distributed over $[0,1]$ according to a distribution F_v. Consider random variables $\{\varepsilon_i\}_{i=1}^{\infty}$ that are distributed i.i.d. conditional on v and strictly affiliated with V; the conditional distribution is given by $f(\varepsilon|v)$ with full support on $[0.1]$. Assume that in addition there is a random variable Z that is independent of both ε_i and V and is distributed according to a log-concave density function f_z with full support on $[0,1]$. Bidders observe private signals $s_i = z + \varepsilon_i$. We argue that:

Proposition *(i) S_i and V are affiliated. (ii) the price in a $k_n + 1$-price auction in which $k_n/n \to \alpha \in (0,1)$ does not converge to the value of the asset as $n \to \infty$. (iii) the expected value of the asset conditional on all the signals does converge to v, that is $E(V|\{s_i\}) - V \to 0$.*

We prove this result in the Appendix; the logic behind it can be seen from the following example:

Example *Suppose that the value of the asset v is distributed uniformly on $[0,1]$. Random variables ε_i are distributed independently conditional on v according to a distribution function $f(\varepsilon_i|v) = 1 - v + 2v\varepsilon$ and Z is distributed uniformly on $[0,1]$. The players obtain signals $s_i = z + \varepsilon_i$.*

In this case the conditional density of signals is:

$$f_s(s|v) = \begin{pmatrix} s(vs + 1 - v) \text{ if } s \in [0,1] \\ (2-s)vs + 1 - v \text{ if } s \in (1,2]) \end{pmatrix}$$

which implies that $\frac{\partial^2}{\partial s \partial v} \ln f_s(s|v) = \frac{1}{(vs+1-v)^2} > 0$. Hence, s and v are strictly affiliated.

Now, note that if all the signals are observed, in the limit it is possible to determine the value of the asset: the lowest ε_i allows determination of z and then finding the median of $s_i - z$ it is possible to calculate the value $v = \frac{1}{2}\frac{2m-1}{m(1-m)}$.

However, consider a uniform price auction, in which half of the players obtain the good and pay $\frac{n}{2} + 1$ bid. In the limit the price is a function of the median signal s, m_s. But for a given median m_s there are in general many combinations of v and z that are consistent with it, even in the limit. To *see this, note that the median m_s is in the range $\left[0, \frac{2+\sqrt{2}}{2}\right]$ and in the limit:*

$$m_s = z + \frac{\sqrt{1+v^2} - (1-v)}{2v}$$

Therefore unless $m_s = 0$ or $\frac{2+\sqrt{2}}{2}$ (which are zero-probability events) the distribution of v conditional on m_s is non-degenerate.

It is interesting to note that while a static auction such as the $k+1$-price auction does not aggregate information, a dynamic auction, such as an English auction does. Results similar to Kremer (2002) imply that the price in an English auction converges to the value of the asset conditional on $n - k_n + 1$ signals. In our case this is more than sufficient to determine the asset's value, as we need only the signal of some quantile and the lowest signal.

Appendix

<u>Proof of Lemma ref: diff:</u>

We first note a few properties of the expected value of the asset conditional on the $l-th$ order statistic being equal to s, $E(V_n|Y_n(l) = s)$, which is given by:

$$\frac{\int vF(S_i = s|v)^{n-l} f(S_i = s|v)(1 - F(S_i = s|v))^{l-1} f(v)dv}{\int F(S_i = s|v)^{n-l} f(S_i = s|v)(1 - F(S_i = s|v))^{l-1} f(v)dv} \qquad \#$$

Because $\lim_{s\uparrow 1} \frac{1-F(s|v)}{(1-s)f(s|v)} = 1$ (by continuity of $f(s|v)$, for details see PS p.1257), taking the limit of the above expression we get:

$$E(V_n|Y_n(l) = 1) = \frac{\int vf(S_i = 1|v)^l f(v)dv}{\int f(S_i = 1|v)^l f(v)dv}$$

Hence, $E(V_n|Y_n(l) = 1)$ does not depend on n.

For $s \in [0,1)$ define $g(s,l) = \frac{\int vf(S_i=s|v)(1-F(S_i=s|v))^{l-1} f(v)dv}{\int f(S_i=s|v)(1-F(S_i=s|v))^{l-1} f(v)dv}$. Note that $g(s,l)$ is continuous in s and by strict affiliation (assumption A2) it is strictly increasing in l. Using again $\lim_{s\uparrow 1} \frac{1-F(s|v)}{(1-s)f(s|v)} = 1$ we can define $g(1,l)$ as the limit as $s \uparrow 1$, which is:

$$g(1,l) \equiv \lim_{s\uparrow 1} g(s,l) = E(V_n|Y_n(l) = 1)$$

Next consider a fixed $\delta \in \left(0, \frac{1}{b}\right)$ and some $s > 1 - \frac{\delta}{n}$. By assumption A3 $F(s|v) > 1 - \frac{b\delta}{n}$ for all v, which implies that:

$$F(s|v)^{n-l} > F(s|v)^n > \left(1 - \frac{b\delta}{n}\right)^n > 1 - b\delta$$

Using (ref: eq1) and $F(s|v)^n - l \in (1 - b\delta, 1)$ we conclude that:

$$E(V_n|Y_n(l) = s) \geq (1 - b\delta)g(s,l)$$

For n is large enough, continuity of $g(s,l)$ and $s > 1 - \frac{\delta}{n}$ implies that $g(s,l) > (1 - b\delta)g(1,l)$. Therefore:

$$E(V_n|Y_n(l) = s) > (1-b\delta)^2 g(1,l)$$

We continue by showing that for any fixed l, $\Pr(Y_n(l) > 1 - \frac{\delta}{n})$ is bounded away from zero. This probability is higher than the probability that exactly l signals are above $1 - \frac{\delta}{n}$. We bound the probability of this event by conditioning on $V = v$.

$$\Pr\left(Y_n(l) > 1 - \frac{\delta}{n}\right) \geq \binom{n}{l} F\left(S_i = 1 - \frac{\delta}{n}|v\right)^{n-l}\left(1 - F\left(S_i = 1 - \frac{\delta}{n}|v\right)\right)^l$$

A3 implies that $1 - F(S_i = 1 - \frac{\delta}{n}|v) > \frac{a\delta}{n}$. As before we note that $F(S_i = 1 - \frac{\delta}{n}|v)^{n-l} > 1 - b\delta$ which provides a bound:

$$\Pr\left(Y_n(l) > 1 - \frac{\delta}{n}\right) \geq (1-b\delta)\binom{n}{l}\left(\frac{a\delta}{n}\right)^l$$

For large enough n we have $\binom{n}{l} > \frac{n^l}{2^l l!}$. Combining, since l is fixed we conclude that for any l there exists some $\alpha > 0$ so that $\Pr(Y_n(l) > 1 - \frac{\delta}{n}) > \alpha$ for all n large enough.

Finally, consider the k-th and $k+1$ order statistics. As we noted above, $g(1,k+1) > g(1,k)$; set δ^* so that $(1-b\delta^*)^2 g(1,k+1) > g(1,k)$. The Lemma then follows as we know that the probability that $Y_n(k+1)$ is above $1 - \frac{\delta^*}{n}$ is bounded away from zero and for any $s > 1 - \frac{\delta^*}{n}$ we have: $E(V_n|Y_n(k+1) = s) > (1-b\delta^*)^2 g(1,k+1) > g(1,k) = E(V_n|Y_n(k) = 1) \geq E(V_n|Y_n(k) \geq s)$.

<u>Proof of Proposition ref: prop1:</u>

(i) Note that

$$f_s(s|v) = f_s(\varepsilon + z|v) = \int f(\varepsilon|v) f_z(s-\varepsilon)d\varepsilon$$

We assume that ε and v are affiliated so $f(\varepsilon|v)$ is log super-modular in ε and v. Also note that $\frac{\partial^2 \log f_z(s-\varepsilon)}{\partial\varepsilon\partial s} = -\frac{\partial^2 \log f_z(s-\varepsilon)}{\partial^2 z}$. Hence, if f_z is log-concave then $f_z(s-\varepsilon)$ is log super-modular in s and ε. One can then apply Lemma 4 and Theorem 1 in Athey (2002) to conclude that S_i are affiliated with V.

(ii) The price is a function of the $k_n + 1$-highest signal or the $\alpha -$ quantile and hence does not contain more information. We argue that knowing the value of the quantile is not sufficient to determine the value of the asset. To see this let $h(v) \equiv F^{-1}_{\varepsilon|v}(\alpha)$ denote the value for which conditional on $V = v$ the probability that ε is lower than $h(v)$ is given by α. Standard results (see Herbert (1981)) imply that the $k_n + 1$-highest ε_i converges to $h(V)$ and hence the $k_n + 1$ highest signal converges to $Z + h(V)$. Conditional on $V = v$ this random variable has full support on $[h(v), 1 + h(v)]$. Since $h(v) \in (0,1)$ we conclude that there is overlap in the support for different values for V and hence the claim follows.

(iii) Since the lowest signal converges to Z, by subtracting it from the $k_n + 1$ highest signal we obtain in the limit the $k_n + 1$ highest ε_i. As noted in (ii) this converges in the limit to $h(v)$. Strict affiliation implies that $h(v)$ is increasing which implies that we are able to determine v.

References

40pt **Athey, S.** (2002), “Monotone Comparative Statics Under Uncertainty,” ***Quarterly Journal of Economics,*** 117, 187-223.

40pt **Herbert, A. David** (1981), “Order Statistics,” ***John Wiley & Sons, second edition***.

40pt **Hong, H. and M. Shum** (2003), “Rates of Information Aggregation in Common Value Auctions,” ***Journal of Economic Theory***, forthcoming.

40pt **Kremer I.** (2002), "Information Aggregation in Common Value Auctions," ***Econometrica***, 70, 1675-1682.

40pt**Milgrom, P. and R. Weber** (1982), "A Theory of Auctions and Competitive Bidding," ***Econometrica,*** 50, 1089–1122.

40pt **Pesendorfer, W. and J. Swinkels** (1997), "The Loser's Curse and Information Aggregation in Common Value Auctions," ***Econometrica,*** 65, 1247–1282.