Computing Reserve Prices and Identifying the Value Distribution in Real-world Auctions with Market Disruptions*

William E. Walsh
CombineNet, Inc.
Fifteen 27th St.
Pittsburgh, PA, USA
wwalsh@combinenet.com

David C. Parkes
SEAS
Harvard University
Cambridge, MA, USA
parkes@eecs.harvard.edu

Tuomas Sandholm
Computer Science Dept.
Carnegie Mellon University
Pittsburgh, PA, USA
sandholm@cs.cmu.edu

Craig Boutilier
Dept. of Computer Science
University of Toronto
Toronto ON, Canada
cebly@cs.toronto.edu

Introduction

Single-good ascending auctions, including the English Auction and its close variants (e.g. eBay), are the most widely used type of auction. Hence, effective strategies for such auctions can have an enormous economic impact. To maximize profit, a seller should try to set a reserve price high enough to extract the highest bidder's full value without blocking the bidder out altogether. In isolated auctions where bidders have static, independent, identically distributed, private values from a known distribution, it is well-understood how to compute the optimal reserve price. However, these assumptions rarely hold in practice. First, the value distributions are not known *ex ante*. Second, auctions rarely run in isolation [2, 7]. Third, value distributions are typically non-stationary.

Estimating the value distribution from bids is non-obvious and non-trivial because the distribution of bids is not the same as the distribution of values [4, 6]. Jiang and Leyton-Brown [6] addressed how "hidden bids" in online auctions can skew the bidding distribution away from the underlying value distribution. They were able to effectively infer the value distribution when given the parametrized form. Haile and Tamer's [4] method infers bounds on the distribution, making no *a priori* assumptions about the bidder valuations and very minimal assumptions about bidder behavior. However, their approach was not complete, as it did not provide a way to choose the reserve price within the bounds. Additionally, neither of the aforementioned approaches addressed the issue of non-isolated auctions or non-stationary distributions. Although Juda and Parkes [7] generalize Haile and Tamer's work somewhat to allow for bidders that participate in multiple auctions, and Gerding et al. [2] analyzed reserve pricing in the presence of multiple competing auctions, those models are simplifications. Indeed, performing a full game-theoretic analysis of setting reserve prices in a dynamic real-world context is prohibitively complex.

We present an automated methodology and system for computing reserve prices for real-world ascending auctions. Our initial system is based on the approach of Haile and Tamer [4], but with the addition of our own technique for computing a specific distribution within their bounds. We adjust the distribution to account for competing sales channels and non-stationarities. We fielded our system in a live two-month trial on real auctions, which demonstrated the effectiveness of our approach, but also revealed its inability to adapt to a drop in bidding coinciding with the appearance of a competing seller. In response, we developed a new Bayesian technique for adjusting to sudden shocks in the market without explicitly having to model their cause. Throughout, we make minimal modeling assumptions and use only that information which is readily available from past auctions.

The Core Method

Consider a standard isolated English auction in which bidders have independent private values for a single good, drawn from the CDF $F(v)$. The seller places value v_0 on the good and specifies a public reserve price of r.

Haile and Tamer [4] showed how to compute upper and lower bounds F_U and F_L, such that, $\forall v$, $F_L(v) \leq F(v) \leq F_U(v)$. They make no assumptions about the form of the distribution and make only minimal assumptions about bidding behavior, namely 1) bidders do not bid more then they are willing to pay, and 2) bidders do not allow an opponent to win at a price they are willing to beat. Unlike previous nonparametric approaches (e.g., [1, 3, 8]), equilibrium behavior is *not* assumed.

The method for computing bounds considers only the highest bids (or only the publicly available *revealed* high bids in eBay-style auctions), and uses information only from auctions with at least two bids. The latter aspect will be of importance when we must estimate the distributions when the past reserve prices were too high.

Under certain regularity conditions [4], the optimal reserve for both English and eBay-style auctions with proxy bidding is given by

$$\arg\max_p (p - v_0)[1 - F(p)]. \quad (1)$$

Although Haile and Tamer described how to compute bounds on the value distribution, they did not describe how to select a particular distribution within the bounds. Our first contribution is a method for estimating a particular distribution $\hat{F}(v)$ within the bounds. Our system then uses $\hat{F}(v)$ in place of $F(v)$ in Eq. 1 to compute a reserve price.

*This work was funded by and performed at CombineNet, inc.

We first compute $F_L(v)$ and $F_U(v)$ for a fine mesh of discrete values of v, according to Haile and Tamer's method. We then pick a particular distribution $\hat{F}(v) = F_L(v) + \alpha(F_U(v) - F_L(v))$, $\alpha \in [0, 1]$ by simulating the revenue and finding a best fit to the actual revenue R.

To estimate the revenue $\hat{R}_\alpha$ for a candidate α, we first compute the empirical distribution $\tilde{N}$ of the number of bidders in an auction, according to the numbers of bidders actually observed. Then we perform a number of simulations, and compute $\hat{R}_\alpha$ as the average revenue from the simulations. For a given simulation, we sample a number n from $\tilde{N}$, and then sample n values v_i from $\hat{F}(v)$. We evaluate the revenue for the simulation as the second highest of the values v_i. In order to avoid the possible distortion of reserve prices, we use data only from auctions with two or more bidders in computing $\tilde{N}$ and R. To determine the correct α, we exploit the fact that $\hat{R}_\alpha$ must be non-decreasing in α, and perform a binary search within $[0, 1]$ until $|\hat{R}_\alpha - R|$ is sufficiently small.

In the full paper we demonstrate that our extension to Haile and Tamer's method can compute good approximations of both $F(v)$ and the optimal reserve price, using data from as few as 25 auctions. The approximations are very close using data from 100 auctions.

A Live Test of Reserve Pricing

We developed an automated reserve pricing system that uses the method described above and tested it in a two-month live trial. During the trial, the system was used by a company that auctions returned goods in bulk to resellers on the Internet. Each auctioned good was a bundle of a variety of electronics in various conditions. 176 auctions were conducted in the trial, producing over $270,000 in revenue.

Modeling the Value Distributions A significant challenge to computing reserve prices was that no two auctions sold the exact same bundle of items. We addressed this with two methods. First, we performed all computations in terms of *recovery*, that is the dollar value of a bid normalized by the bundle's wholesale cost. Historically, all bundles sold for a recovery of less than 1, allowing us to bound the search for $\hat{F}(v)$ and the reserve price. Second, following Haile and Tamer [4], we identified various attributes that affected price and weighted past data according to a product of Gaussian kernels [5] on the attributes.

A nonlinear regression showed a strong seasonality effect on auction prices, with higher prices in the months before Christmas (when people buy for the holidays) and lower prices in January and February (when people return goods after the holidays). Thus we used day of the year as the attribute in one of the Gaussian kernels.

Since we did not have useful detailed data on the composition of each bundle, we identified easy-to-compute proxies that signaled the relative value of different bundles. Nonlinear regressions showed that the final prices tended to increase with the the total wholesale cost of the bundle and also with the average wholesale cost of individual items in the bundle. We used both of these measures as attributes in the product of Gaussian kernels.

Another issue was that the company ran multiple auctions simultaneously, and the goods in simultaneous auctions could be considered (partial) substitutes to some of the bidders. Since a complete model of valuations and the optimal reserve price would be prohibitively complex, we worked with the isolated auction model, but with the additional narrow assumption that the number of simultaneous auctions would tend to affect the revenue in the auction. In fact, a regression showed that revenue tended to decrease with the number of simultaneous auctions. Thus we used the number of simultaneous auctions as the attribute in a Gaussian kernel.[1]

Note that we did not encode the actual trend of valuations with respect to the attributes. That is, the system did not know the actual effects on value due to seasonality, cost, average cost, or number of simultaneous auctions *a priori*. Rather, when computing the reserve price for a good with particular attributes, the system simply weighted the data from a past auction more or less heavily based on the similarity of the past auction attributes to the current one (according to a product of Gaussian kernels). The actual effect of the attributes on the value distribution was automatically inferred by the system.

The Auction Setup The auctions were eBay-style auctions—that is English auctions with fixed end dates and proxy bidding. On average, 6.5 auctions were run at the same time. Before our trial, the company always set the reserve price to a recovery of 0.1 (i.e., 10% of the wholesale price), which was clearly suboptimal since they were able to sell the goods for an average of 0.18 through a separate, fixed-price channel. Because the company had historically set the reserve price so low, the auctions had always received bids above the reserve. Since our system would compute significantly higher reserve prices we had to contend with unsold goods. We decided with the company that, if a good didn't sell at auction, it would be reposted one more time to auction at the same reserve price. If it did not sell in that second auction either, it would be combined with less desirable items into a much larger bundle and sold through a fixed-price channel to one of a small set of select buyers (in order to minimize the effect that a bidder in the auction would feel that he could buy the same bundle through a secondary channel if the item did not sell in the auction). Since goods could be sold through the fixed-price channel for an average recovery of 0.18, we took that to be our v_0.

Since we allowed unsold goods to be reposted, a higher reserve price than specified by Eq. 1 would be optimal. We first computed an initial value using Eq. 1, then searched for

[1]The existence of simultaneous auctions for substitute goods does not actually decrease a bidders' underlying *value* for the goods. Rather, it will tend to decrease the resulting *bids* for the goods because there are multiple options. However, rather than modeling the substitutabilities and the resulting effect on bidding, we instead pretend as if the presence of simultaneous auctions actually changes the underlying values for an individual good, and that bidding proceeds as if the auction is run in isolation. Although this is admittedly an approximation, the success of our live trial bears out its effectiveness.

the optimum at higher prices. To estimate the revenue that would be obtained at a candidate reserve r, we performed simulations in the same manner as we did to compute $\hat{F}(v)$, but with the reserve set to r and with the auction rerun one more time, with a new set of bidders, if the reserve was not met. We used one year of past auction data as the historical data, and we did not update that data set during the trial. We performed 2000 simulations each in the computation of $\hat{F}(v)$ and the repost-adjusted reserve price. For each Gaussian kernel, we used a bandwidth of 0.3 times the standard deviation of the attribute. In evaluating the effectiveness of our system, we compared the recovery obtained using reserve prices from our system with the recovery obtained during the same time period the previous year.[2]

Results from the Trial Our system was quite effective during the first month of the trial. The recovery increased by 5.6%, from 0.303 to 0.32 (statistically significant at a 99.5% confidence level), as compared to the same period during the previous year.[3] This improvement includes the price of goods sold at auction as well as unsold goods later sold at a fixed price averaging 0.18. The reserve price varied, with different attributes of the goods, in the range $[0.26, 0.33]$. The recovery of goods successfully sold at auction was in the range $[0.285, 0.433]$ and 10% of the auctions did not receive bids exceeding the reserve. The experimental results also agreed well with our simulation, which suggested a recovery improvement of 4.8% and that 9.2% would not receive bids above the reserve.

But an unpleasant surprise occurred during the second month of the trial: the performance of our system dropped dramatically, resulting in a decrease in recovery of 19%, as compared to the previous year. During this period, 76% of the auctions did not receive bids exceeding the reserve price. Investigations by the company revealed that a competing seller entered the market during the second month, and we believe that the increased competition resulted in the drop in bid prices. To test this, we turned off the reserve pricing system after the second month and gathered data for another month. The recovery during the third month decreased by 13% as compared to the same period during the previous year. We conclude that the 19% decrease in the second month was due largely to entry of the competitor, with some portion of the decrease due to a reserve price that was too high in light of the new market environment.

Adapting to Market Disruptions

Our deployed system failed to adapt to the changing market conditions in the live trial because it did not incorporate the data from auctions run during the trial (as explained above, it only used the historical data set as the basis for setting reserve prices). A straightforward solution would have been to simply include auction results generated during the trial, perhaps with a heavier weighting on more recent data. Although this would allow our system to adapt if bids had increased, it could not not have been effective for decreasing bids. The problem is that the core system makes use of data only from auctions with at least two bids. When the reserve price is too high, as was most surely the case in the second month, the only data available is from the highest value bidders—when auctions receive any bids at all. As a result, when bids decrease, the estimate of $\hat{F}(v)$, and hence the reserve price, is increasingly skewed upward. If our system had incorporated the new auction data in a straightforward fashion, it would likely have performed worse.

We needed a method for inferring something about the part of the distribution that is hidden by the reserve price.[4]

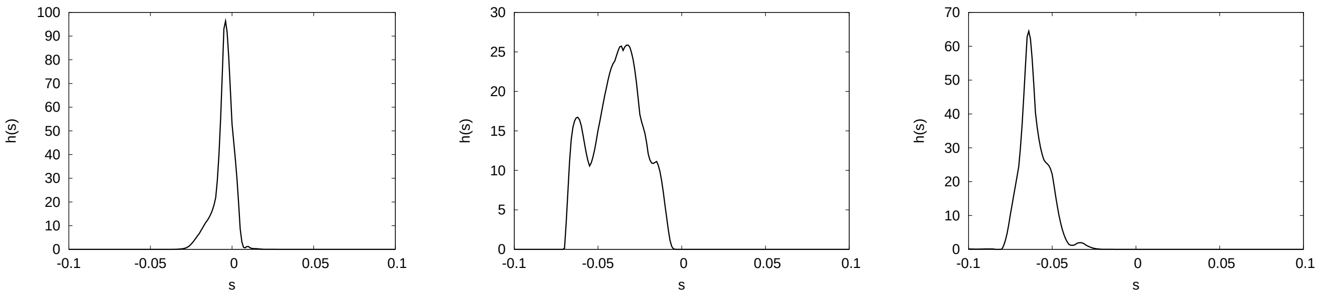
Modeling Market Changes Changes in observed bidding behavior can potentially arise from a variety of underlying causes. Because of the complexity of modeling multiple effects in the world market, we chose a simplified, unified model whereby we assume changes in bidding reflect fundamental changes in the value distribution. Our approach is to model changes in bidding as a *shift* in the underlying value distribution. That is, a drop in bidding is indicative of a leftward shift of the value distribution while an increase in bidding is indicative of a rightward shift. To estimate the amount of shift, we look at a window of w previous auctions, and assume that a shift may have occurred immediately before those auctions were conducted. Using the predicted probability of an auction receiving no bids above the reserve, and the actual history of auctions receiving no bids, we compute the posterior distribution of the amount of shift in the value distribution. We then shift the value distribution by the estimated amount and compute the reserve price from the shifted distribution.

Computing the Distribution of Shift Assume for now that all auctions have the same attributes. Let $\hat{F}(v)$ be our current estimate of the cumulative density function of values, as determined with our initial method (i.e., not taking into account any shifting). We assume that the true distribution $F(v)$ is equal to $\hat{F}_s(v)$, the distribution $\hat{F}(v)$ shifted by $s \in \Re$. We use $s > 0$ to indicate a rightward shift, and $s < 0$ to indicate a leftward shift. Let $\delta(s)$ be the event that the shift is s and let $g(s)$ be the prior probability density function of $\delta(s)$. Let r_i be the reserve price for past auction $i \in w$, and let $\pi_{i,s}$ be the probability that i would receive no bids above r_i, given s. For a given s, it is straightforward to determine $\hat{F}_s(v)$, and, as indicated earlier, we can compute $\pi_{i,s}$ via simulation using $\hat{F}_{i,s}(v)$ and the observed distribution of the number of bidders. Let $\gamma_{i,s}$ indicate the probability of the actual outcome of auction i, given r_i and s. That is, if i received at least one bid, then $\gamma_{i,s} = 1 - \pi_{i,s}$,

[2] We compared recovery, rather than actual dollar revenue, because the wholesale cost of the goods was not the same during both periods.

[3] This amount of improvement was good considering that the returned electronics auctions had historically been quite competitive. We would expect an even greater improvement for other types of goods that tend to get fewer bidders.

[4] Jiang and Leyton-Brown [6] addressed the problem of how to infer the distribution when low-value bidders do not bid because the auction price has already exceeded their value (because high-value bidders bid first). However, their method, like our first method, ignores the potentially valuable information available from auctions that receive no bids at all.



(a) (b) (c)

Figure 1: Distribution $h(s)$ of shift s in the trial as collected: (a) at the end of the first month, (b) in the middle of the second month, (c) at the end of the second month.

otherwise $\gamma_{i,s} = \pi_{i,s}$. To compute the posterior distribution $h(s)$ of s, we use Bayes' rule:

$$h(s) = \frac{\Pi_{i \in w} \gamma_{i,s} g(s)}{\int_{-\infty}^{\infty} \Pi_{i \in w} \gamma_{i,s} g(s) ds}. \quad (2)$$

It is not immediately clear how to estimate $g(s)$, given that it is a prior on the error of our estimate. One possibility would be to substitute the prior distribution over the amount of shift that occurs in the actual market during any given short period of time. For the present work, we simply assume an uninformative prior and let the data fully determine $h(s)$. To compute $h(s)$, we first compute each $\gamma_{i,s}$ for discrete values of s within a reasonable range. In our real-world trial, we normalized values into $[0, 1]$, so we can safely assume that $s \in [-1, 1]$. Then, for any desired value of $s \in [-1, 1]$, we compute $h(s)$ numerically using Romberg integration and polynomial interpolation on $\gamma_{i,s}$ [9].

Our goal is to estimate the optimal reserve price for a new auction. We can compute $\hat{F}(v)$ using the initial method and $h(s)$ given the Bayesian method above. Ideally, we would compute the reserve r^* that maximizes the expected revenue. A less expensive approach would be to compute from $h(s)$ the expected shift $\bar{s}$, shift $\hat{F}(v)$ by $\bar{s}$, and then compute the optimal reserve price from the shifted distribution. This would give a particularly good approximation when $h(s)$ is unimodal with small variance. If auctions have different attributes, then we assume that the goods are similar enough that the amount of the shift is the same for all goods. For each past auction i we compute a separate $\hat{F}_{i,s}(v)$ based on its attributes, and use that to determine $\gamma_{i,s}$. Then we compute $\hat{F}(v)$ for the good we are auctioning presently, shift it by $\bar{s}$, and compute the reserve price.

Adaptation Experiment To demonstrate our adaptation method, we applied it to the data obtained in our live test. At three points in time during the trial, we computed the unshifted value distribution $\hat{F}(v)$ (using attributes from a representative auction run at that time) and the optimal reserve, r, given $\hat{F}(v)$. Using a window of 30 auctions we computed $h(s)$, $\bar{s}$, $\hat{F}_{\bar{s}}(v)$ (the distribution $\hat{F}(v)$ shifted by $\bar{s}$), and the optimal reserve price $r_{\bar{s}}$ for $\hat{F}_{\bar{s}}(v)$. We compute a product of Gaussian kernels using attributes as in the live test. Unlike in our live test, we now use all data obtained up to a given date. Figs. 1 (a)–(c) show that $h(s)$ does indicate a shift in the distribution during the second month of the trial. Furthermore $r_{\bar{s}}$ decreases from r in response to the shift. At the end of the first month, $r = 0.300$, $\bar{s} = -0.005$, and $r_{\bar{s}} = 0.305$. In the middle of the second month, $r = 0.250$, $\bar{s} = -0.04$, and $r_{\bar{s}} = 0.235$. Finally, at the end of the second month, $r = 0.285$, $\bar{s} = -0.061$, and $r_{\bar{s}} = 0.225$. (Note that r is different at each point because the auction attributes are different.)

We cannot evaluate how accurate $r_{\bar{s}}$ is from the live data because we do not know the true value distribution. In the full paper, we investigate the interaction of the augmented system with simulated auctions and market changes. The results show that our approach can effectively adapt and compute an accurate reserve price after a market shift.

Conclusions

Automated reserve pricing can be an effective approach to increasing revenue in auctions for a variety of industries. Our methodology provides a scalable way to compute appropriate reserve prices in complex, real-world settings, while using only inputs that are readily available in practice.

References

[1] Susan Athey and Philip A. Haile. Identification of standard auction models. *Econometrica*, 70:2107–2104, 2002.

[2] Enrico H. Gerding, Alex Rogers, Rajdeep K. Dash, and Nicholas R. Jennings. Sellers competing for buyers in online markets: Reserve prices, shill bids, and auction fees. In *IJCAI*, 2007.

[3] Emmanuel Guerre, Isabelle M. Perrigne, and Quang Vuong. Optimal nonparametric estimation of first-price auctions. *Econometrica*, 68:525–574, 2000.

[4] Philip A. Haile and Elie Tamer. Inference with an incomplete model of English auctions. *J. of Political Economy*, 111(1):1–51, 2003.

[5] W. Härdle. *Applied Nonparametric Regression*. Cambridge University Press, 1990.

[6] Albert Xin Jiang and Kevin Leyton-Brown. Bidding agents for online auctions with hidden bids. *Machine Learning J.*, 67(1-2):117–143, 2007.

[7] Adam I. Juda and David C. Parkes. The sequential auction problem on eBay: An empirical analysis and a solution. In *ACM EC*, pages 180–189, 2006.

[8] Tong Li, Isabelle M. Perrigne, and Quang Vuong. Structural estimation of the afficliated private value auction model with an application to ocs auctions. *Rand J. of Economics*, 33:171–193, 2002.

[9] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C: Second Edition*. Cambridge University Press, 1992.