





Performance Evaluation 65 (2008) 10-31



www.elsevier.com/locate/peva

# Performance of the Vickrey auction for digital goods under various bid distributions

M. Naldi\*, G. D'Acquisto

Università di Roma "Tor Vergata", Dip. di Informatica Sistemi Produzione (DISP), Via del Politecnico 1, 00133 Roma, Italy

Received 1 October 2005; received in revised form 26 July 2006; accepted 3 February 2007 Available online 8 February 2007

#### Abstract

The generalized Vickrey auction is analysed as a mechanism for setting prices of digital goods. A selection of models of the bidders' behaviour (embodied by the probability distribution of their bids' values) are employed to derive the optimal auctioneer's choice as to the number of items to sell. The satisfaction levels of both the auctioneer and the bidders are taken into account through the evaluation of the expected revenues, the percentage of winning bidders and their aggregated utilities under each of the probabilistic scenarios considered.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Auctions; Vickrey; Digital goods; Bid distributions

## 1. Introduction

Auctions are considered an efficient mechanism to set prices, since they require little a priori knowledge of the potential market's response to the offer of the items for sale. Since the value of the item to the bidders is signalled by the bids themselves (when the auction mechanism is correctly defined in order to get a so-called truthful auction), auctions represent an implementation of value pricing: the final price is consistent with the winning bidders' valuations (meaning that it is lower or equal to them). In addition, the auctioneer, gathering the whole set of bids, has an immediate picture of the present state of the demand function, which is what every a priori mechanism for setting prices would require. As a result, auctions are increasingly employed on e-commerce platforms.

Though the mass image of auctions is typically linked to collection items, they have recently gained acceptance in the context of the information and communication technology, where their use has been envisaged for several applications:

- the electronic commerce of networking goods [1], where the latter term may include anything from a leased line to a single communication session, e.g. the transmission of data packets over a communication network, where a bid is attached to each packet [2,3] or to each bandwidth request [4];
- the assignment of radio spectrum to mobile operators (see e.g. the assignment of UMTS licences in Europe), where the number of participants may be very low [5–7];
- the sale of hosting services for websites [8].

<sup>\*</sup> Corresponding author. Tel.: +39 06 72597269; fax: +39 06 72597532.

E-mail addresses: naldi@disp.uniroma2.it (M. Naldi), giuseppe.dacquisto@inwind.it (G. D'Acquisto).

The field we focus on in this paper is the application of auctions to the electronic commerce of digital goods (e.g. music, books, and any information or entertainment item that can be exchanged digitally), which is naturally aimed at a mass market (see e.g. [9]).

Since digital goods have negligible marginal costs of production (the cost of copying is practically limited to the physical support), they can be considered to be in unlimited supply. An added feature of auctions, when dealing with such goods, is therefore that the seller can freely choose the quantity of items to sell (i.e. the number of copies of the original item to distribute), his/her only aim being the maximization of revenues. Both deterministic and randomized mechanisms for this choice have been proposed and evaluated in [10,11], considering the revenues as the only performance metric and adopting a worst case approach, so to provide bound guarantees on the revenues attained under the proposed mechanism with respect to a benchmark. Though correct, this approach leaves a wide uncertainty on the actual performance of the auction, since the range of the achieved revenues framed by the bounds is often very large. As a matter of fact, this analysis is accompanied in [10] by an evaluation of the expected revenues conducted by simulation under several probability models for the bids' distribution.

The aim of this paper is to investigate, by analytical means in addition to simulation, the performance of auctions for digital goods. As in [10] a number of probability models, representative of different attitudes of the bidders with respect to the value of the item for sale, are considered for the distribution of bids' values. In this paper we limit ourselves to the analysis of a deterministic mechanism, the Vickrey auction [12], which assigns the item to the k most generous bidders at the price offered by the k+1-st bidder (i.e. the largest among the rejected bids). This mechanism, hereafter named as the single-price auction since the price is equal for all the winning bidders, has the advantage of being perceived as fair and satisfies the important condition of truthfulness, i.e. of inducing the bidders to declare the true value of the item (which varies from bidder to bidder). Implementations of the Vickrey auction have been proposed in [2,3]. As a benchmark, against which we compare the single price mechanism, we use the differential pricing scheme, where all users are satisfied, each one at a price equal to his/her bid; the revenues associated to the differential pricing scheme are the largest possible under an auction mechanism (this condition makes it suitable as a benchmark). Though in this latter case each user pays for the value he attributes to the good on auction, this scheme is legitimate only if the auctioneer is not compelled to make the price setting mechanism public; if that's not the case, each bidder, knowing in advance that he/she will win the item anyway, would bid the lowest possible value; hence, the differential pricing scheme revenues are unattainable in practice.

For each bids distribution model the optimal choice of the number of sold items is derived; this is all the auctioneer needs to state its revenue-maximizing strategy. We are here envisaging an a priori implementation of the Vickrey mechanism, where the revenue optimization is a priori, i.e. it can be operationally achieved if the parameters of the bids' distribution are known or at least accurately estimated (which can be accomplished on-line, during the bids' arrival process), so that the auctioneer doesn't need to compute the revenues for all values of k after the auction closure, in order to pick up the value maximizing the revenues. This is opposed to an a posteriori optimization, where the optimal k is chosen after the N bids are all known. As a further benchmark we also consider here the revenues obtained in the a posteriori case.

The single price mechanism is evaluated through a set of metrics. The satisfaction of the seller is assessed through the expected revenues resulting from the optimal choice of the number of sold items, compared with both benchmarks, represented by the expected revenues under the above mentioned differential pricing and the a posteriori mechanism. We consider two additional metrics, to keep into account the bidders' satisfaction, namely the percentage of successful bidders and, since the final price in a Vickrey auction is lower than the value of the accepted bids, the savings achieved by them. These metrics are evaluated for all the probability models by making use of the theory of order statistics.

Since the a priori optimization relies on the knowledge of the bids distributions, its quality may depend critically on the assumptions made to this regard. It is therefore relevant to examine what happens in the case of mismatching, i.e. when the distribution adopted for the purpose of selecting the number of items to sell is different from the actual one. The performance is here examined under all the possible mismatching situations.

The organization of the paper is as follows. In Section 2 the auction mechanism is thoroughly described, together with the evaluation procedure, including the evaluation metrics. Then, Sections 3 through 7 are devoted to the application of the procedure described in Section 2 to each of the probabilistic scenarios considered. The analysis of mismatching is reported in Section 8, while he comparison of the a priori vs. the a posteriori implementation of the Vickrey mechanism is conducted in Section 9.

## 2. Evaluation framework

In this section we set the basic framework employed to evaluate the auction mechanisms recalled in the Introduction.

The bidders send their bids to the auctioneer over the network, within a given time framework, after which the auction is closed and the auctioneer declares both the setting price and the winning bidders. Hereafter we consider the roles of the auctioneer and of the seller, though generally distinct, to be coincident, without loss of generality. This online auction is of the sealed bids kind, so that each bid is known to the bidder and to the auctioneer only. Given the potentially worldwide scale of an online auction, it is assumed that the number N of bids, though random, may be very large. Each bid is considered to be the realization of the random variable X obeying the probability density function  $f_X(x)$ . Alternatively the N bids can be seen as the realizations of N random variables  $X_i$  (i = 1, 2, ..., N), independent and identically distributed as X. In addition, the bids can be ordered in a non-increasing fashion  $x_{(1)} \ge x_{(2)} \ge \cdots \ge x_{(N)}$ , so to be considered as the realizations of N random variables (order statistics)  $X_{(1)}, X_{(2)}, \ldots, X_{(N)}$ .

Here we consider five scenarios for the bids' distribution, represented by the following probability models: uniform, triangular, Gaussian, exponential and Pareto. The first three of them are adequate to model the case of bids more or less clustered around a central value, with the uniform and the triangular having a finite support and therefore allowing for the definition of a minimum bid, while the Gaussian model, though admitting negative values, is here used for sake of completeness and for a range of its moments such that negative values are extremely unlikely. The exponential and the Pareto models represent instead more scattered sets of bids, but again with allowance for a minimum bid. This is a wider choice than that adopted in [10], where the uniform model and a discrete version of the Pareto model (Zipf's law) were considered in a similar analysis conducted by simulation. Buyers' behaviour is modelled similarly in [13], but with a slightly more restricted choice of models, since four buyers vs. price curves are considered to mimic the buyers' valuation of the item for sale in the context of dynamic pricing for perishable goods. In particular, on that simulator the curve where the number of buyers decreases as the price increases is analogous with our exponential and Pareto distributions, and the curve with mid-peaking has the same role as our triangular model.

The metrics employed to evaluate the single price auction are:

- the expected revenues under an a priori choice of the number of accepted bids;
- the ratio between the aforementioned revenues and those associated to the differential pricing, hereafter named the efficiency;
- the revenue loss due to the a priori choice of the number of accepted bids instead of the a posteriori one, hereafter named the a priori loss;
- the percentage of bids that are accepted, hereafter named the Satisfied Bidders Ratio (SBR);
- the ratio between the savings obtained by the winning bidders and the sum of their bids, hereafter named the Satisfied Utility Margin (SUM).

The maximization of the expected revenues is considered to be a goal of the auctioneer, so that the first three metrics are indicators of the goodness of the auction to the auctioneer.

However, auctions are a two-sided deal; the two other metrics are therefore proxies for the level of satisfaction of the bidders. In fact, the bidder considers his/her bid as the correct value of the good, and will judge as acceptable any price equal to (as in the differential pricing case) or below (as in the single price case) it, but will undoubtedly be frustrated if that bid is discarded (and therefore discouraged to take part to further auctions). Hence an appropriate measure of satisfaction is the percentage of accepted bids: the more winners are declared the more the auction participants are satisfied; measuring this type of satisfaction is the aim of the SBR indicator.

Instead, if we focus on the winning bidders, their level of satisfaction will be increased if the price requested of them is below their bid, so that they achieve savings with respect to their bid; the SUM indicator measures how large these savings are.

Employing metrics for the auctioneer's satisfaction as well as for the bidders' satisfaction is in our view necessary to reach an equilibrate evaluation of the auction mechanism, though often neglected in the literature on auctions.

We now define the general expressions for the above introduced metrics in the two auction mechanisms under comparison.

If we accept the r largest bids, setting a price equal to the highest declined bid (i.e. the r + 1st), as in the r-items Vickrey auction which is the chosen mechanism for the single price case, the revenues are

$$R = rX_{(r+1)},\tag{1}$$

and are random under any choice of the number of sold items. If the latter is k, the expected revenues are

$$E_k[R] = kE[X_{(k+1)}].$$
 (2)

These revenues can be maximized by a proper choice of the number of accepted bids, i.e. of k. Here we define the single price strategy as that using the optimal k that maximizes the expected revenues:

$$E[R_{\rm sp}] = \max_{k} (kE[X_{(k+1)}]) = k_{\rm opt} E[X_{(k_{\rm opt}+1)}]. \tag{3}$$

Given the choice of the Vickrey auction, setting the number of items to sell is the most important decision for the auctioneer; hence, solving expr. (3) for k produces the auctioneer's strategy.

In this context the optimization is a priori, i.e. it can be operationally achieved if the parameters of the bids' distribution are known or at least accurately estimated (which can be accomplished on-line, during the bids' arrival process), so that the auctioneer doesn't need to compute the revenues for all values of k after the auction closure, in order to pick up the value maximizing the revenues. The evaluation of a single price strategy with a posteriori optimization, where the optimal k is chosen after the N bids are all known, so that

$$k_{\text{opt}} = k : R = \max_{k} k x_{(k+1)},$$
 (4)

is used as a benchmark through the evaluation of the a priori loss, defined as

$$APL = \frac{\max_{k} k x_{(k+1)} - E[R_{sp}]}{\max_{k} k x_{(k+1)}},$$
(5)

whose expected values will be evaluated in this paper.

The evaluation of (2) or (3) requires the knowledge of the expected value of the order statistics, whose general expression is [14]:

$$f_{X_{(r)}}(x) = \frac{N!}{(r-1)!(N-r)!} f_X(x) [1 - F_X(x)]^{r-1} [F_X(x)]^{N-r} \quad r = 1, 2, \dots, N.$$
 (6)

Instead, the expected revenues in the differential price strategy, where all the bids are accepted, each one at its own value, are:

$$E[R_{\rm dp}] = \sum_{i=1}^{N} E[X_{(i)}] = \sum_{i=1}^{N} E[X_i] = N \cdot E[X]. \tag{7}$$

Of course the revenues under the differential pricing strategy are always larger than in the single price case; actually they provide the largest harvest that could be obtained in a sealed bid auction. So, they represent a reference against which the single price strategy can be compared. We can therefore define the efficiency of the single-price auction as

$$Eff = \frac{E[R_{sp}]}{E[R_{dp}]}.$$
(8)

Turning to the metrics concerning the satisfaction of bidders, we can see that the SBR is always 1 for the differential pricing case, so that a comparison can be conducted by simply computing the SBR for the single price case, given by:

$$SBR = \frac{k_{\text{opt}}}{N}; \tag{9}$$

the larger the SBR (bounded by 1) the more satisfied the bidders. In the single price auction we have at most

$$SBR = 1 - \frac{1}{N}.$$
 (10)

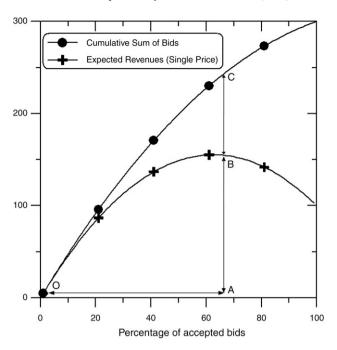


Fig. 1. Metrics for the bidders' satisfaction.

Instead the SUM index is always zero for the differential pricing, while it is equal to:

$$SUM = \frac{\sum_{i=1}^{k_{\text{opt}}} E[X_{(i)} - X_{(k_{\text{opt}}+1)}]}{\sum_{i=1}^{k_{\text{opt}}} E[X_{(i)}]} = 1 - \frac{k_{\text{opt}} E[X_{(k_{\text{opt}}+1)}]}{\sum_{i=1}^{k_{\text{opt}}} E[X_{(i)}]} = 1 - \frac{E[R_{\text{sp}}]}{\sum_{i=1}^{k_{\text{opt}}} E[X_{(i)}]}$$

$$(11)$$

for the a priori single price case. This indicator grows with the level of satisfaction (hence, the differential pricing is the negative reference in this case) and is again bounded by 1. The form (11) suggests that this indicator can also be interpreted as the fraction of revenues that are lost by adopting a single price mechanism instead of a differential pricing one that sells the same number of items.

In this regard it is to be noted that the bidders' utility is often defined (e.g. in [11]) as the difference between the bids' values and the final price; under this definition the SUM index is therefore an aggregated measure of the bidders' utility.

A pictorial view of the metrics used to assess the bidders' satisfaction is shown in Fig. 1, where the segment OA represents the SBR index and the ratio of the segments BC/AC gives the SUM index.

## 3. The uniform distribution of bids

In the uniform model the bids are distributed according to the probability density function

$$f_X(x) = \begin{cases} \frac{1}{h-l} & l \le x \le h \\ 0 & \text{elsewhere,} \end{cases}$$
 (12)

where l and h are respectively the lowest and the highest bid.

In settings modelled by this function the bidders have a fairly common understanding of the range of values of the good for sale, but their individual evaluations may be quite different within that range.

The evaluation of the revenues under the single price strategy requires the computation of the expected value of the revenues and their maximization, i.e. the optimal choice of the number of accepted bids. A straightforward application

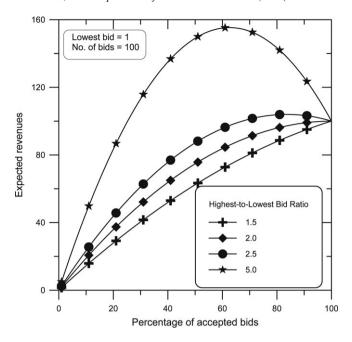


Fig. 2. Influence of the number of accepted bids on the revenues in the uniform case.

of (6) gives us the probability density function of the order statistics in the uniform case:

$$f_{X_{(r)}}(x) = \frac{N!}{(r-1)!(N-r)!} \cdot \frac{(h-x)^{r-1}(x-l)^{N-r}}{(h-l)^N} \quad l \le x \le h$$
(13)

and the expected value of the order statistics

$$E[X_{(r)}] = l + (h - l)\frac{N + 1 - r}{N + 1},\tag{14}$$

so that the expected revenues, when k items are sold, are

$$E_k[R] = kE[X_{(k+1)}] = kl \left[ \left( \frac{h}{l} - 1 \right) \left( 1 - \frac{k+1}{N+1} \right) + 1 \right].$$
 (15)

As expected, the revenues grow linearly with the lowest bid and with the highest-to-lowest bid ratio. When seen as a function of the number of accepted bids, expression (15) is a quadratic form, so that we can expect an optimal value of k to exist. Actually, as can be seen in the sample case of Fig. 2, the relationship between the revenues and the number of accepted bids can instead be non-monotonic.

In order to determine the precise behaviour of the revenues to this regard we compute the marginal revenues, i.e. the variation in the expected revenues deriving from selling one item more:

$$\Delta_k = E_{k+1}[R] - E_k[R] = l + (h - l) \left( 1 - 2\frac{k}{N+1} \right). \tag{16}$$

The marginal revenues are a decreasing function of the number of sold items and can turn negative, at which point it is not convenient to sell more items. The optimal value of k is the largest one for which the incremental revenues are positive, i.e.  $\Delta k > 0$ . By a simple manipulation of (16) we obtain

$$k_{\text{opt}} = \left| \frac{N+1}{2} \cdot \frac{h/l}{h/l-1} \right|,\tag{17}$$

where the symbol  $\lfloor z \rfloor$  denotes the floor function, returning the largest integer less or equal to z. So the optimal number of objects to sell can be determined once the number of bidders and the highest-to-lowest bid ratio are known.

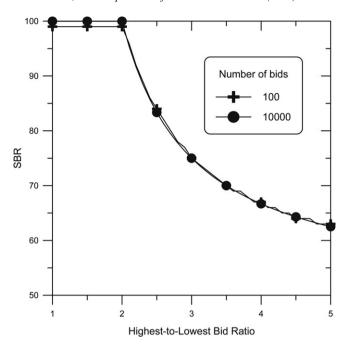


Fig. 3. Optimal percentage of accepted bids in the uniform case.

However, the expression (17) is meaningful only if the resulting  $k_{\text{opt}}$  is not larger than N-1, or, in other terms, if the following condition is satisfied:

$$\frac{N+1}{2} \cdot \frac{h/l}{h/l-1} < N. \tag{18}$$

This condition entails a constraint on the highest-to-lowest bid ratio h/l:

$$h/l > \frac{2N}{N-1} \approx 2,\tag{19}$$

confirming what the visual inspection of the sample case of Fig. 2 had suggested.

In the end, therefore, all bids, excepting the lowest one, should be accepted if the h/l ratio is lower than 2; if that's not the case, the expression (17) has to be applied. As can be seen in Fig. 3, the percentage of bids to be accepted declines as the h/l ratio exceeds 2 and depends negligibly on the number of bids. The limiting value, when the bids are widely scattered (i.e. when the h/l ratio tends to infinity), is half of the submitted bids. Therefore, when the uniform model applies, a large fraction of the bidders are satisfied.

The resulting revenues are approximately

$$E[R_{\rm sp}] \approx \begin{cases} Nl + h - l & h/l < 2\\ \frac{N}{4} \cdot \frac{h^2}{h - l} & h/l \ge 2. \end{cases}$$
 (20)

The other indicator of the bidders' satisfaction, the SUM index, can be readily computed by replacing the expected revenues in the general expression (11). The result is practically constant SUM  $\approx 1/3$ , meaning that on the average the winning bidders pay just two-thirds of their valuations.

We now turn to the differential pricing strategy. In this case the expected revenues are quite simply given by

$$E[R_{\rm dp}] = N \cdot E[X] = N \frac{h+l}{2},\tag{21}$$

and grow linearly with both the number of bids and the average bid value.

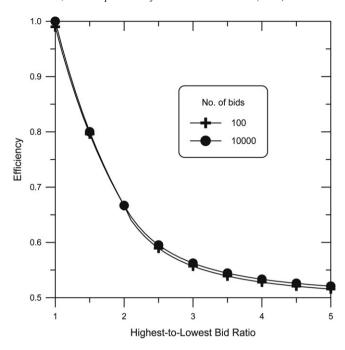


Fig. 4. Efficiency of the single price strategy in the uniform case.

We are now in a position to compare the two strategies as to their expected revenues, by using the efficiency as defined by (8). By forming the ratio of the two previous expressions we have:

Eff = 
$$\frac{E[R_{\rm sp}]}{E[R_{\rm dp}]} \approx \frac{2}{h/l+1} \left( 1 + \frac{h/l-1}{N} \right) \quad h/l < 2,$$
 (22)

$$Eff = \frac{E[R_{sp}]}{E[R_{dp}]} \approx \frac{1}{2[1 - (l/h)^2]} \quad h/l \ge 2.$$
 (23)

In both cases we see in Fig. 4 that the efficiency is practically independent of the auction size N, but decreases very fast as the bids' dispersion grows; the asymptotic limit is  $\frac{1}{2}$ , meaning that the single price auction's revenue are halved with respect to what the differential pricing auction could provide.

# 4. The triangular distribution of bids

The second distribution we consider is the triangular one (also known as the Simpson model), whose probability density function is limited, as the uniform distribution, to within the [l; h] interval:

$$f_X(x) = \begin{cases} 4\frac{x-l}{(h-l)^2} & l \le x < \frac{l+h}{2} \\ 4\frac{h-x}{(h-l)^2} & \frac{l+h}{2} \le x \le h \\ 0 & \text{elsewhere.} \end{cases}$$
 (24)

This model applies when there is a significant convergence of the bidders' valuations towards a common value (the central value of the [l; h] range), but some random fluctuations around this are present.

The straightforward application of (6) gives us the p.d.f. of the order statistics:

$$f_{X_{(r)}}(x) = \begin{cases} \frac{N!}{(r-1)!(N-r)!} 2^{2+N-r} \frac{(x-l)^{2(N-r)+1}}{(h-l)^{2N}} [(h-l)^2 - 2(x-l)^2]^{r-1} & l \le x < \frac{l+h}{2} \\ \frac{N!}{(r-1)!(N-r)!} 2^{r+1} \frac{(h-x)^{2r-1}}{(h-l)^{2N}} [(h-l)^2 - 2(h-x)^2]^{N-r} & \frac{l+h}{2} \le x \le h. \end{cases}$$
(25)

If we sell *k* items the expected revenues are therefore:

$$E_{k}[R] = \frac{N!}{(k-1)!(N-k-1)!} \cdot \frac{1}{(h-l)^{2N}} \left\{ 2^{N+1-k} \int_{l}^{(l+h)/2} x(x-l)^{2(N-k)-1} [(h-l)^{2} - 2(x-l)^{2}]^{k} dx + 2^{k+2} \int_{(l+h)/2}^{h} x(h-x)^{2k-3} [(h-l)^{2} - 2(h-x)^{2}]^{N-k-1} dx \right\}.$$
(26)

Though the two integrals present in the foregoing equation can be solved by parts, the resulting expression is quite cumbersome. It is therefore more enlightening to look at some of the expected revenues curves reported in Fig. 5, obtained under the same conditions (as to the number of bids and to their extreme values) of the uniform case depicted in Fig. 2. The overall behaviour is very similar to that obtained in the uniform case, the main differences being that now the peak value is slightly larger and is always reached when not all the bids are accepted.

The optimal percentage of accepted bids, reported in Fig. 6, has therefore an initial sharp decrease (contrasting with the nearly complete acceptance of the uniform case) but then falls more gently, staying well above the 50% limit (even nearly 80% when the highest bid is as high as 5 times the lowest bid). In addition, it is practically independent of the number of bids.

Since the differential pricing benchmark is given again by expr. (21), the efficiency has the same trend as in the uniform case, depending negligibly on the number of bids and decreasing when the dispersion of bids increases. When the h/l ratio is 5 the efficiency has reached 60%, slightly above its asymptotic value of 50% (see Fig. 7).

Turning again to the satisfaction of the winning bidders, the Satisfied Utility Margin, depicted in Fig. 8, is instead a growing function of the dispersion of bids; it depends negligibly on the number of bids and tends to the same limiting value 1/3 of the uniform case. The clustering of bids around a common value has therefore a negative effect on the bidders' utilities.

## 5. The Gaussian distribution of bids

This model represents again the case where the bids are scattered around a central value, embodied by the expected value  $\mu$  of the bids. The probability density function of the bids is therefore

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$
 (27)

The degree of scattering around the central value is described by the standard deviation  $\sigma$ . This model would allow for negative values (absent in real auctions) but not for a minimum bid (present in the uniform and triangular value). In order to reconcile these two shortcomings we consider (a)  $(\mu, \sigma)$  couples such that the probability of having negative values is extremely low; (b) a conventional definition of the range of values (theoretically infinite in the Gaussian model) as six-times the standard deviation (resembling the 6-sigma concept of quality control) so that the probability of having values outside this range is again very low. The second condition can be expressed as the following relationship between the standard deviation of the Gaussian model and the conventional values for the highest and the lowest bid

$$h - l = 6\sigma. (28)$$

For the analysis of the Gaussian model, unlike the previous two cases (uniform and triangular), no exact general expressions exist for the order statistics. The approximations proposed in [14,19] fail when applied to the whole range of orders as in our case. In the Gaussian case we therefore resort to simulation for the evaluation the metrics of interest, by generating 10 000 replications for each evaluation.

The resulting behaviour is not dissimilar from those seen in the previous two sections. The expected revenues reported in Fig. 9 grow with the number of accepted bids up to a maximum which is attained when nearly all the bids are accepted (at least for the lowest h/l ratios). More precisely the optimal number of accepted bids decreases gently as the dispersion grows.

In fact the percentage of satisfied bidders keeps larger than that encountered in the uniform or triangular models (see Fig. 10), being nearly 85% even when the highest bid is five-times larger than the lowest one, and, as shown by the simulation, this property doesn't seem to depend on the number of bidders.

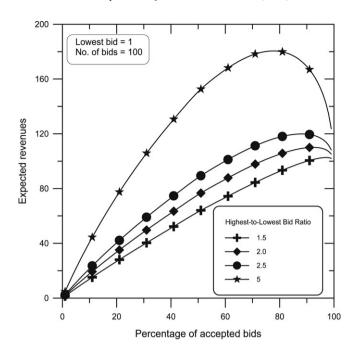


Fig. 5. Influence of the number of accepted bids on the revenues in the triangular case.

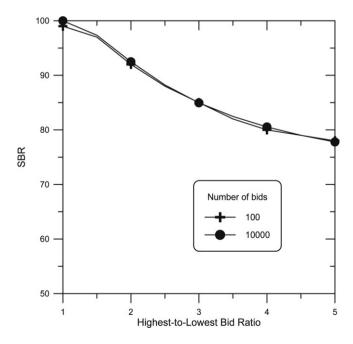


Fig. 6. Optimal percentage of accepted bids in the triangular case.

The efficiency also shows the same trend as the previous models: independence of the dimension of the auction and lowering as the bids dispersion grows (see Fig. 11).

Finally the savings obtained by the winning bidders grow as the dispersion grows, but with an asymptotic limit slightly lower than that achieved in the triangular case (see Fig. 12).

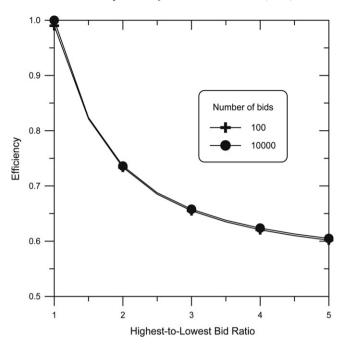


Fig. 7. Efficiency of the single price strategy in the triangular case.

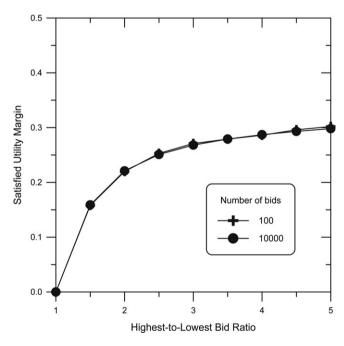


Fig. 8. Satisfied utility margin for the triangular distribution.

# 6. The exponential distribution of bids

According to the exponential model the bids follow the probability density function:

$$f_X(x) = \lambda e^{-\lambda(x-c)} \quad x \ge c. \tag{29}$$

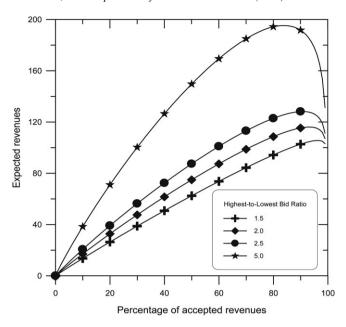


Fig. 9. Influence of the number of accepted bids on the revenues in the Gaussian case.

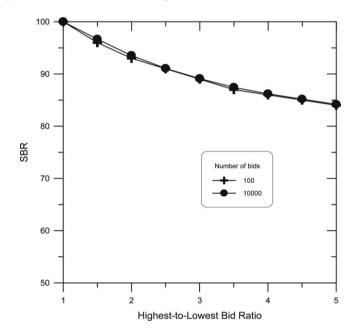


Fig. 10. Optimal percentage of accepted bids in the Gaussian case.

This model allows for a minimum bid value, represented by the parameter c (the standard case with c=0 corresponds instead to the case where any bid larger than zero is possible); its expected value is  $c+1/\lambda$  and its standard deviation is  $1/\lambda$ . In contrast to the previous two models, now the bids are widely scattered, with low valuations being commoner than large ones, the more the larger  $\lambda$  is.

The expected value of the order statistics, that we need to compute the expected revenues and optimize them, can be derived from [15]:

$$E[X_{(r)}] = c + \frac{1}{\lambda} \sum_{i=r}^{N} \frac{1}{i}.$$
(30)

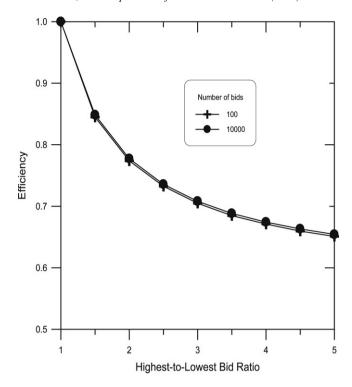


Fig. 11. Efficiency of the single price strategy in the Gaussian case.

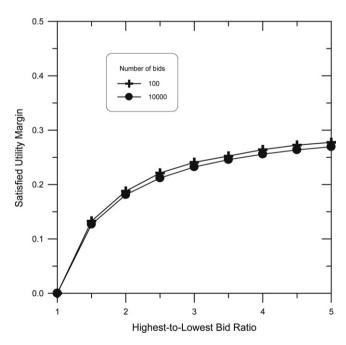


Fig. 12. Satisfied utility margin for the Gaussian distribution.

For the generic choice of the number of sold items the expected revenues are therefore

$$E_k[R] = kE[X_{(k+1)}] = \frac{k}{\lambda} \sum_{i=k+1}^{N} \frac{1}{i} + kc.$$
(31)

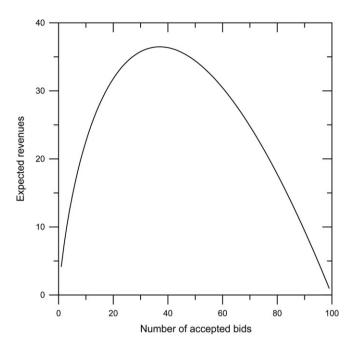


Fig. 13. Influence of the number of accepted bids on the revenues in the exponential case (N = 100).

In the standard case, since the expected revenues result linearly dependent on the quantity  $1/\lambda$ , we can gain some insight on their behaviour by simply plotting them for the case  $\lambda = 1$ . This is done in Fig. 13 for the case of N = 100 bids, where a well defined peak is present; a similar behaviour is expected for all values of N.

In the general case the optimal k can be recovered by forming again the marginal revenues:

$$\Delta_{k} = E_{k}[R] - E_{k-1}[R] > 0$$

$$\Delta_{k} = \frac{k}{\lambda} \sum_{i=k+1}^{N} \frac{1}{i} + kc - \left(\frac{k-1}{\lambda} \sum_{i=k}^{N} \frac{1}{i} + (k-1)c\right) = \frac{k}{\lambda} \left(\sum_{i=k+1}^{N} \frac{1}{i} - \sum_{i=k}^{N} \frac{1}{i}\right) + \frac{1}{\lambda} \sum_{i=k}^{N} \frac{1}{i} + c$$

$$= -\frac{k}{\lambda} \cdot \frac{1}{k} + \frac{1}{\lambda} \sum_{i=k}^{N} \frac{1}{i} + c = \frac{1}{\lambda} \left(\sum_{i=k}^{N} \frac{1}{i} - 1\right) + c > 0 \Rightarrow \sum_{i=k}^{N} \frac{1}{i} - 1 > -c\lambda \Rightarrow \sum_{i=k}^{N} \frac{1}{i} > 1 - c\lambda \tag{32}$$

and solving for the largest k such that the inequality (32) is satisfied.

This is certainly true if  $c\lambda > 1$ , since the right-hand term becomes negative; this condition amounts to the following constraint on the expected value of the bids:

$$c\frac{1}{E[X]-c} > 1 \Rightarrow c > E[X]-c \Rightarrow E[X] < 2c \Rightarrow c < E[X] < 2c. \tag{33}$$

In this special case the optimal number of bids is N-1.

If that's not the case, i.e.  $c\lambda < 1$ , the exact determination of the optimal number of bids from the inequality (32) is always possible, but doesn't provide a closed form expression. Instead, if we consider the following approximation of the harmonic series [16]:

$$\sum_{i=1}^{N} \frac{1}{i} \approx \gamma + \ln(N) + \frac{1}{2N},\tag{34}$$

where  $\gamma$  is Euler's constant, the inequality  $\Delta_k > 0$  can be written as

$$\gamma + \ln(N) + \frac{1}{2N} - \left(\gamma + \ln(k - 1) + \frac{1}{2(k - 1)}\right) > 1 - c\lambda$$

$$\Rightarrow \ln(N) + \frac{1}{2N} - \ln(k - 1) - \frac{1}{2(k - 1)} > 1 - c\lambda$$

$$\Rightarrow \ln(k - 1) + \frac{1}{2(k - 1)} < \ln(N) + \frac{1}{2N} + c\lambda - 1.$$
(35)

Neglecting the terms (much smaller than unity) 1/2N and 1/2(k-1) the optimal number of sold items results from the following passages:

$$\ln(k-1) < \ln(N) + c\lambda - 1$$

$$\Rightarrow k - 1 < N \cdot e^{c\lambda - 1}$$

$$\Rightarrow k_{\text{opt}} = \lfloor 1 + N \cdot e^{c\lambda - 1} \rfloor.$$
(36)

The standard case c = 0 has the simpler expression

$$k_{\text{opt}} = \left\lfloor \frac{N}{e} + 1 \right\rfloor. \tag{37}$$

These solutions are quite accurate: for 10 000 bidders the optimal number of accepted bids supplied by (37) is 3677 against the correct value of 3679; for 100 bidders expr. (37) gives the exact solution  $k_{\text{opt}} = 37$ . Hence, the optimal percentage of accepted bids is roughly:

$$SBR \approx \begin{cases} 1/e & c = 0 \\ e^{c\lambda - 1} & c\lambda < 1 \\ 1 & c\lambda \ge 1. \end{cases}$$
(38)

In sharp contrast with the uniform and the triangular cases, when the standard exponential model applies we can get the largest revenues by rejecting a percentage of the bids as large as roughly 65%, a constant and large fraction of the total number. However, the optimal percentage of rejections decreases to zero as the minimum value of the bids grows.

For the standard case the resulting optimal expected revenues are:

$$E[R_{\rm sp}] = \left\lfloor \frac{N}{e} + 1 \right\rfloor \frac{1}{\lambda} \sum_{i=k+1}^{N} \frac{1}{i} \approx \frac{N}{e} \cdot \frac{1}{\lambda} \left[ \gamma + \ln(N) + \frac{1}{2N} - \gamma - \ln\left(\frac{N}{e}\right) - \frac{1}{2N/e} \right] \approx \frac{N}{\lambda e}. \tag{39}$$

Turning to the general case of non-zero minimum bid, a similar expression holds for the  $c\lambda < 1$  case:

$$E[R_{\rm sp}] \approx \frac{N}{e} e^{c\lambda} \left( \frac{1}{\lambda} + c \right) = \frac{N}{e} e^{c\lambda} E[X],$$
 (40)

while the  $c\lambda \ge 1$  condition leads to

$$E[R_{\rm sp}] = \frac{N-1}{\lambda} \cdot \frac{1}{N} + (N-1)c = (N-1)\left(\frac{1}{N\lambda} + c\right). \tag{41}$$

Though slightly more than one-third of the bidders are satisfied in the standard case (leading to a low value of the SBR index of bidders' satisfaction), the winners can boast a larger utility margin than in the previous cases. In fact, the SUM index for the standard case is:

$$SUM = 1 - \frac{\frac{k_{\text{opt}}}{\lambda} \sum_{i=k_{\text{opt}}+1}^{N} \frac{1}{i}}{\frac{1}{\lambda} \sum_{i=1}^{k_{\text{opt}}} \sum_{j=1}^{N} \frac{1}{j}} = 1 - \frac{1}{1 + \frac{1}{\sum_{i=k_{\text{opt}}+1}^{N} \frac{1}{i}}} \approx 1 - \frac{1}{1 + \frac{1}{\ln(N) - \ln(k_{\text{opt}}) + \frac{1}{2N} - \frac{1}{2k_{\text{opt}}}}} \approx \frac{1}{2},$$
(42)

so that the winning bidders pay on the average just half their bids values.

If we now turn to the measure of satisfaction for the auctioneers, the differential pricing benchmark is

$$E[R_{\rm dp}] = N \cdot E[X] = N\left(c + \frac{1}{\lambda}\right),\tag{43}$$

And therefore the efficiency is approximately:

$$\operatorname{Eff} = \frac{E[R_{\rm sp}]}{E[R_{\rm dp}]} \approx \begin{cases} \frac{1/e \quad c = 0}{\frac{N}{e}} e^{c\lambda} \left(c + \frac{1}{\lambda}\right)}{N\left(c + \frac{1}{\lambda}\right)} = e^{c\lambda - 1} \quad c\lambda < 1\\ \frac{c + 1/N\lambda}{c + 1/\lambda} \quad c\lambda \ge 1. \end{cases}$$
(44)

We note that for the standard case the efficiency reduces to the constant value  $\text{Eff} \approx 1/e$ , quite smaller than the asymptotic lower bound encountered in the uniform and the triangular cases. For the standard case therefore the more scattered distribution of bids' values with clustering around lower values, represented by the exponential model, is detrimental to the auctioneers' satisfaction, irrespective of the expected value of bids.

## 7. The Pareto distribution of bids

The last model we consider in this paper for the distribution of bids is the Pareto one, whose probability density function is

$$f_X(x) = -\frac{\alpha}{c} \left(\frac{x}{c}\right)^{-(\alpha+1)} \quad x > c. \tag{45}$$

In this scenario bids are always larger than the lower bound c. Lower values are commoner than large ones, with the value of the shape parameter  $\alpha$  dictating the relative weight of lower bids, but very large valuations are possible: lower values of  $\alpha$  correspond to fatter tail of the pdf, but if  $\alpha < 1$  not even the expected value of X is finite. If  $\alpha > 1$  we have

$$E[X] = \frac{\alpha}{\alpha - 1}c. \tag{46}$$

For this distribution the expected value of the r-th order statistic is [17,18]:

$$E[X_{(r)}] = c \frac{N!}{(r-1)!} \cdot \frac{\Gamma\left(r - \frac{1}{\alpha}\right)}{\Gamma\left(N + 1 - \frac{1}{\alpha}\right)},\tag{47}$$

where  $\Gamma(\cdot)$  is the Gamma function.

By exploiting the recurrent relation

$$\Gamma(z+1) = z\Gamma(z),\tag{48}$$

we obtain the expected revenues when the auctioneer decides to sell k items:

$$E_k[R] = kE[X_{(k+1)}] = kc\frac{N!}{k!} \frac{\Gamma\left(k+1-\frac{1}{\alpha}\right)}{\Gamma\left(k+1-\frac{1}{\alpha}\right)\prod\limits_{i=k+1}^{N}\left(i-\frac{1}{\alpha}\right)} = kc\frac{1}{\prod\limits_{i=k+1}^{N}\left(1-\frac{1}{\alpha}\cdot\frac{1}{i}\right)}.$$

$$(49)$$

Here the lowest possible bid c doesn't play any role in the choice of the optimal number of sold items. Instead, by looking at some plots of expr. (49) for different values of the shape parameter, reported in Fig. 14, we note that all the curves are monotonic.

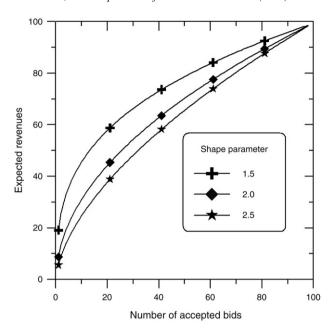


Fig. 14. Number of accepted bids on the revenues in the Pareto case (c = 1, N = 100).

In fact, this is a general property, that can be shown by computing the variation of revenues when the auctioneer sells k + 1 items instead of k:

$$\frac{E_{k+1}[R]}{E_k[R]} = \frac{k+1}{k} \cdot \frac{k!}{(k+1)!} \cdot \frac{\Gamma\left(k+2-\frac{1}{\alpha}\right)}{\Gamma\left(k+1-\frac{1}{\alpha}\right)} = 1 + \frac{1-\frac{1}{\alpha}}{k}.$$
 (50)

Since  $\alpha > 1$  the above ratio is always larger than 1, so that the optimal number of accepted bids is always N - 1. Under the single price mechanism we have therefore

$$E[R_{\rm sp}] = c(N-1) \approx cN,\tag{51}$$

and, since

$$E[R_{\rm dp}] = NE[X] = N\frac{\alpha}{\alpha - 1}c,\tag{52}$$

the efficiency is

$$Eff = \frac{c(N-1)}{N\frac{\alpha}{\alpha-1}c} \approx 1 - \frac{1}{\alpha}.$$
 (53)

The efficiency is therefore larger the more concentrated the bids are near the lower bound c, but can be extremely small if the shape parameter is slightly larger than 1.

Since the number of winning bidders is N-1, we expect their utilities to be very large. Actually we have:

$$SUM = 1 - \frac{E[R_{sp}]}{\sum_{i=1}^{N-1} E[X_{(i)}]} = 1 - \frac{E[R_{sp}]}{\left(\sum_{i=1}^{N} E[X]\right) - E[X_{(N)}]} \approx 1 - \frac{cN}{N\frac{\alpha}{\alpha - 1}c - c} \approx \frac{1}{\alpha} = 1 - \frac{c}{E[X]};$$
 (54)

hence, the savings of the bidders are larger the larger their bids.

# 8. Mismatched choices of the probability model

The analysis has been conducted so far by assuming that the distribution is known in advance or at least can be inferred on the fly. Actually some mistake can enter the picture so that the distribution assumed for the decision as to the number of bids to accept can be different from the actual one. It is quite relevant to examine what happens if this mismatching occurs, namely how much the expected revenues suffer from the wrong choice.

For this purpose we divide the distributions of interest into two classes, the first one being composed of the three ones exhibiting a distribution of bids around a central value (i.e. the uniform, the triangular and the Gaussian), while the second class gathers the exponential and the Pareto model, where the probability decreases monotonically as the bid value grows. The reason for this subdivision is that it is quite reasonable to assume that the general trends of the distributions belonging to the two classes are so different that it is unlikely to mistake one class for another. We therefore limit the examination to the cases of mismatching between distributions belonging to the same class. In addition, for the comparison to be correct, we assume that the parameters of the two distributions being mistaken are computed on the basis of the same values for the first two moments (expected value and variance).

As to the first class we are therefore to examine all the possible mismatching combinations, which are 6, by considering all the couples formed by the actual distribution and the wrongly assumed one. We compute the two moments corresponding to the characteristic values of the actual distribution, and then derive the characteristic values of the wrongly assumed distribution corresponding to the two moments previously computed.

We start with the two cases resulting when the actual distribution is uniform. In this case the moments corresponding to the distribution characteristic values (the l and h values) are:

$$\mu = \frac{l+h}{2}$$

$$\sigma = \frac{h/l}{\sqrt{12}} = \frac{2\mu}{\sqrt{12}} \cdot \frac{h/l-1}{h/l+1}.$$
(55)

These are the moments employed to generate the corresponding Gaussian distribution, while for the triangular distribution the relation between its characteristic values and the two moments is

$$\left. \frac{h}{l} \right|_{t} = \frac{\sigma\sqrt{6} + \mu}{\mu - \sigma\sqrt{6}}.\tag{56}$$

Under the condition of equal moments we evaluate the revenue loss due to the mismatching. The evaluation is conducted by simulation, with 10 000 replications of an auction with 100 participants. The resulting loss is reported in Fig. 15. As can be seen the loss is really low, achieving a maximum value of about 3%.

In both cases the uniform assumption leads to accepting a larger number of bids for the lowest values of the h/l ratio and a lower number otherwise.

We proceed in the same way in the other two cases, i.e. when the actual distribution is triangular or Gaussian. The resulting losses are reported in Fig. 16 and Fig. 17 respectively.

The uniform assumption appears to be much riskier when the actual distribution is triangular or Gaussian than the reverse case

As to the class including the exponential and the Pareto model, the Pareto model has both the first two moments finite if  $\alpha > 2$ , which amounts to the c < E[X] < 2c condition. This is the same constraint we met in the exponential model, leading us to take on all the bids excepting the lowest one, which is exactly the same choice adopted under the Pareto model. Therefore the equality condition between the moments of the two distributions leads to the same choice as to the number of bids to accept and results in no mismatching losses.

# 9. A priori and a posteriori implementations of the Vickrey auction

Deciding a priori what is the optimal number of bids to accept is less efficient than deciding after the bids are all known. This loss of efficiency, which we name a priori loss, can be quantified by the loss in revenues. In order to evaluate this loss we have simulated again 10 000 replications of an auction with 100 participants for all the five distribution models. In the case of the symmetric probability distributions (uniform, triangular and Gaussian) we have examined how the loss varies as a function of the highest-to-lowest bid ratio, which is the characteristic parameter for

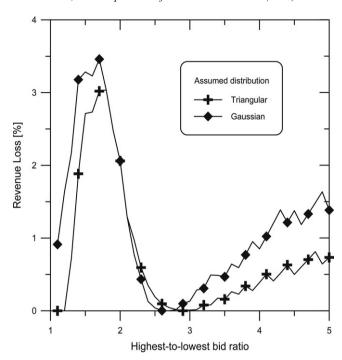


Fig. 15. Revenue loss when the actual distribution is uniform.

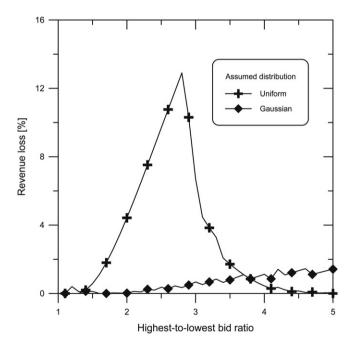


Fig. 16. Revenue loss when the actual distribution is triangular.

these models. The a priori loss, pictured in Fig. 18, is always lower than 4%, though it exhibits an increasing trend as the h/l ratio grows.

For the exponential and Pareto model we use the standard deviation as a characteristic parameter for the bid dispersion, examining, in the exponential case, both the cases of minimum bid equal to zero (standard exponential model) and different than zero (set to 20 in our simulation, as for the Pareto model). The number of replications for

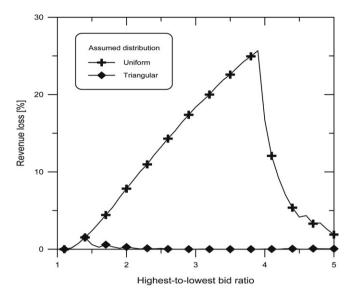


Fig. 17. Revenue loss when the actual distribution is Gaussian.

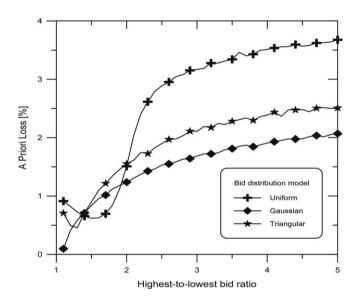


Fig. 18. A priori loss for the symmetrical probability models.

each auction of 100 participants is again 10 000. The resulting a priori loss is reported in Fig. 19. Here the loss appears a bit larger for the standard exponential case (though always rather limited), but is instead very low for the general exponential and Pareto models, though it shows a relevant increasing trend as the standard deviation grows in the exponential case.

### 10. Conclusions

For each of the probabilistic scenarios for the distribution of bids the performance parameters of the Vickrey auction (namely the efficiency, pertaining to the satisfaction of the auctioneer, and the percentage of accepted bids plus the aggregated utilities of the winning bidders, pertaining to the satisfaction of the bidders themselves) have been evaluated analytically. In addition, the expression for the optimal choice of the number of items to sell has been obtained, so that the auctioneer can form an appropriate strategy under different scenarios.

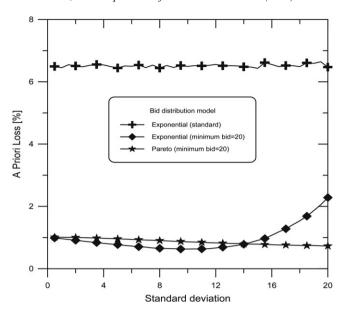


Fig. 19. A priori loss for the exponential and Pareto models.

The three models applying when the bidders' evaluations of the item's value are more or less concentrated around a central value, i.e. the uniform and the triangular one, lead the auctioneer to satisfy the majority of the bidders. As the bidders converge towards a common value, both the percentage of satisfied bidders and the efficiency achieved by the auctioneer grow (of course the absolute revenues are instead larger the larger the expected value of the bids). Both quantities are nevertheless larger than 50%.

On the contrary, for the two scenarios characterized by a prevalence of lower values (though the pdf's tail can be quite fat), i.e. the exponential and the Pareto one, we have contrasting results. The exponential hypothesis leads the auctioneer to accept a constant fraction (1/e) of the bids in the worst case (standard exponential), while, if the Pareto model applies, accepting all the bids is the best decision. The efficiency is anyway very low (1/e) for the standard exponential model and can be even smaller for the Pareto case. If the bids are exponentially distributed, but with a non-zero lower bound the efficiency grows with the bid lower bound up to values very near to the unity.

The mismatching effects are negligible if the distribution of choice is triangular or Gaussian but can be very large if we blindly choose the uniform model. No mismatching occurs if we choose either the Pareto or the exponential model (with equal first and second moments).

The envisaged a priori strategy performs very well in comparison to the traditional a posteriori one, since the a priori loss is limited to some percentage point. The most significant problems are encountered in the case of the exponential distribution, since the loss appears to grow quite fast with the dispersion of the bids.

In all cases the performances are nearly invariant to the auction size (i.e. the number of participants), and hence the strategy indications so far obtained scale very well even for a mass market.

## References

- [1] D. Messerschmitt, J.-P. Hubaux, Opportunities for electronic commerce in networking, IEEE Communications Magazine (1999) 95–98.
- [2] J.K. MacKie-Mason, H.R. Varian, Pricing the internet, in: Proceedings of the Conference on Public Access to the Internet, JFK School of Government, 26–27 May 1993.
- [3] P. Reichl, G. Fankhauser, B. Stiller, Auction models for multi-provider internet connections, in: MMB'99 Workshop, Trier, 22–24 September 1999.
- [4] N. Semret, R.R.F. Liao, A.T. Campbell, A.A. Lazar, Pricing, provisioning and peering: Dynamic market for differentiated internet services and implications for network interconnections, IEEE Journal on Selected Areas in Communications 18 (12) (2000) 2499–2513.
- [5] N.J. Colmenares, The FCC on personal wireless, IEEE Spectrum 31 (5) (1994) 39–46.
- [6] B.Z. Cobbs, Telecommunications, IEEE Spectrum 32 (1) (1995) 30–34.
- [7] W. Webb, P. Marks, Pricing the ether, IEE Review 42 (2) (1996) 57-60.
- [8] M. Bichler, J. Kalagnanam, K. Katircioglu, A.J. King, R.D. Lawrence, H.S. Lee, G.Y. Lin, Y. Lu, Applications of flexible pricing in business-to-business electronic commerce, IBM Systems Journal 41 (2) (2002) 287–302.

- [9] E.H. Durfee, et al., The dynamics of the UMDL service market society, in: M. Klusch, G. Weiss (Eds.), Cooperative Information Agents II, Springer, 1998, pp. 55–78.
- [10] A. Goldberg, J. Hartline, A. Wright, Competitive auctions and digital goods, Intertrust StarLab Technical Report STAR-TR-99-01, April 2001. Also published in the Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms, SODA, 2001.
- [11] Z. Bar-Yossef, K. Hildrum, F. Wu, Incentive-compatible online auctions for digital goods, University of Berkeley SIMS296A-5/CS294-6 Class Project Report, 15 May 2001. Available at: http://www.cs.jhu.edu/~scheideler/club/spring\_02/ziv-baryossef.ps.
- [12] W. Vickrey, Counterspeculation, auctions, and competitive sealed tenders, Journal of Finance 16 (1961) 8-37.
- [13] J. Morris, A simulation-based approach to dynamic pricing, MS Thesis in Media Arts & Science, Massachusetts Institute of Technology, May 2001.
- [14] H.A. David, Order Statistics, J. Wiley, 1981.
- [15] B. Epstein, M. Sobel, Life testing, Journal of the American Statistical Association 48 (263) (1953) 486–502.
- [16] I.S. Gradshteyn, I.M. Ryzhik, Table of Integrals, Series, and Products, Academic Press, 1980. expr. (0.131).
- [17] H.J. Malik, Exact moments of order statistics from the Pareto distribution, Skandinavisk Aktuarietskrift 53 (1-2) (1970) 6-9.
- [18] G. Kulldorf, K. Vannman, Estimation of the location and scale parameters of a Pareto distribution by linear functions of order statistics, Journal of the American Statistical Association 68 (341) (1973) 218–227.
- [19] H. Cramer, Mathematical Methods of Statistics, Princeton University Press, 1946.



Maurizio Naldi graduated cum laude in 1988 in Electronic Engineering at the University of Palermo and then received his Ph.D. in Telecommunications Engineering from the University of Rome "Tor Vergata". After graduation he pursued an industrial career, first at Selenia as a radar designer (1989–1991), and then in the Network Planning Departments of Italcable (1991–1994), Telecom Italia (1995–1998), and WIND (1998–2000) where he was appointed Head, Traffic Forecasting & Network Cost Evaluation Group. In the 1992–2000 period he was active in the standardization bodies (ETSI and ITU), in particular as Associate Rapporteur for Broadband Traffic Measurements and Models at ITU Study Group 2. Since 2000 he is with the University of Rome at Tor Vergata, where he is now Aggregate Professor.



**Giuseppe D'Acquisto** graduated cum laude in 1995 in Electronic Engineering and received the PhD in Telecommunications in 1999 from the University of Palermo-Italy, with a thesis on rare event simulation. After PhD he started a consulting career, working for Telco Operators and ICT companies, with a focus on market forecasting, traffic engineering and cost accounting. He collaborates with the Universities of Rome Tor Vergata and Palermo in researches in the area of simulations and stochastic optimization. He is the author of more than 20 publicatons on these topics.