The expected payoff to Internet auctions

Laurens de Haan · Casper G. de Vries · Chen Zhou

Abstract In an Internet auction, the expected payoff acts as a benchmark of the reasonableness of the price that is paid for the purchased item. Since the number of potential bidders is not observable, the expected payoff is difficult to estimate accurately. We approach this problem by considering the bids as a record and 2-record sequence of the potential bidder's valuation and using the Extreme Value Theory models to model the tail distribution of the bidder's valuation and study the expected payoff. Along the discussions for three different cases regarding the extreme value index γ, we show that the observed payoff does not act as an accurate estimation of the expected payoff in all the cases except a subclass of the case $\gamma = 0$. Within this subclass and under a second order condition, the observed payoff consistently converges to the expected payoff and the corresponding asymptotic normality holds.

Keywords Internet auction · Extreme value theory · Record sequence

L. de Haan
University of Tilburg, Tilburg, The Netherlands

L. de Haan
University of Lisbon, Lisbon, Portugal

L. de Haan · C. G. de Vries · C. Zhou
Erasmus University Rotterdam, Rotterdam, The Netherlands

C. G. de Vries
Tinbergen Institute, Rotterdam, The Netherlands

C. Zhou (✉)
Economics and Research Division, De Nederlandsche Bank,
P.O. Box 98, 1000AB Amsterdam, The Netherlands
e-mail: zhou@few.eur.nl

∅ Springer

AMS 2000 Subject Classifications Primary—91B26; Secondary—62G32

1 Introduction

Internet auctions (IA) provide an easily accessible platform for trade. This has increased the extent of the market for many items to nationwide or even global markets. IA improve the matching between the supply and demand. Bulow and Klemperer (1996) showed that under the hypothesis of the Independent Private Values Paradigm (IPVP), the seller is better off if the market is larger. By the same fact, buyers are also better off, since it becomes more likely that the agent with highest consumer surplus is matched with the seller.

Auction theory derives the optimal bid functions for specific auction mechanisms, such as the Dutch (descending price) or the English (ascending price) auctions, and given a specific demand function. The demand function is modeled as a distribution of valuations of the object to be auctioned. Both the seller and the buyer have an interest in knowing the final price that might materialize to answer such questions as: Is it worthwhile to put the item up for sale? and, is it worth my time to bid? Sellers may want to extrapolate from a single auction to predict total revenues from repeat sales.

Under IPVP, the expected price is predicted as follows. For any standard auction, the Revenue Equivalence Principle (REP) holds, which means that under IPVP, the expected revenue of the seller does not depend on the auction mechanism, see e.g. Krishna (2002).[1] The expected revenue for a standard mechanism is equal to the expectation of the second highest order statistic of the valuations. For the Vickrey or second-price sealed bid auction, this is easily shown to be the case. In a second-price sealed bid auction, the winner pays the second highest bid. Since for this auction agents have an incentive to exactly bid their valuations, the claim follows. Auction theory shows that this revenue result holds for all standard auctions.

Most IAs have two mechanisms for placing a bid, i.e. the manual bid and the proxy bid. This induces a hybrid of an English auction and a Vickrey auction. In an English auction, bidders publicly compete with each other by placing ascending bids. The Vickrey mechanism uses sealed bids, i.e. bidders do not see the bids of competitors. Since both of these two mechanisms are standard, the IA is also standard in the sense of revenue equivalence. Compared to studies of classical auction mechanisms, the empirical analysis of an IA is severely hampered due to the unobserved number of potential bidders. In an IA, besides the *active bidders* who indeed place a bid on the websites, there are also a large number of *potential bidders* who only check the website with or without placing a bid. The number of potential bidders plays a role similar to the number of bidders in classical auction mechanisms, i.e. the number of

---

[1]In classical auction theory, an auction is called *standard* if the rule dictates that the person who bids the highest amount is awarded the object.

bidders sitting in the auction hall. However, for IA, it is hard to observe this number. Firstly, most of the large IA sites do not provide the number of page views. Secondly, even the number of page views would not be a clean estimate of the number of potential bidders, since a potential bidder may check the website multiple times.

Lacking the knowledge of the number of potential bidders severely restricts the statistical analysis of the expected revenue for the seller. Therefore, a few papers recently focus on analyzing data under some additional assumptions regarding the number of potential bidders. Bajari and Hortaçsu (2003) and Paarsch (1992) do not require knowledge of the number of potential bidders, but they assume that the observed bidders are the only potential bidders willing to pay the reserve price. This assumption appears implausible for IA. For instance, de Haan et al. (2008) argued that the actual extent of the IA market, i.e. the number of potential bidders, is far beyond the observed number of active bidders. Alternative approaches are based on modeling the number of potential bidders. For example, McAfee and McMillan (1987) analyzed the case when the number of bidders is stochastic. Another example is Laffont et al. (1995), who assumed that the unknown number of potential bidders is the same across all auctions under consideration.

Still different is Song (2004) who considered the nonparametric estimation of the distribution of bidder's valuation without having any information on the number of potential bidders. Song (2004) argued that without knowing the number of potential bidders, the distribution of the bidder's valuation is not identified if only the payoff, i.e. the second highest order statistic, is observed. But if one can observe the bid history, then using the two top order statistics identifies the parent distribution, even if the number of potential bidders is unknown.

Under the IPVP, the potential bidder's valuations are assumed to be identically and independently distributed random variables. Because the payoff of an IA is the second largest valuation among all the potential bidders,[2] when the number of potential bidders is sufficiently large, the payoff only depends on the tail of the distribution of the bidders' valuations. Thus there is only a need to model the tail of the distribution. Semi-parametric Extreme Value Theory (EVT) provides an approximation to the tail of the distribution. Caserta and de Vries (2005) applied the EVT approach to investigate the expected payoff. However, the number of potential bidders is a major difficulty for their analysis as they assume that the number of actual bids equals the number of potential bidders.

Existing econometric analysis of auctions, c.f. Paarsch (1992), often proceeds on the basis that the number of bidders is known and that the different auctions are homogeneous, possibly controlled for covariates. This allows for

---

[2]In fact, the final payoff should be the second largest valuation plus a minimum increment because the winner has to overbid the second largest valuation. We assume that the minimum increment is negligible compared to the value.

the pooling of the data from different auctions in order to estimate the demand curve (distribution of valuations) and to test for the IPVP. In the current analysis, we do not entertain this maintained homogeneity assumption, as we want to investigate the expected price of a particular, possibly unique, auction. For this purpose the EVT approach appears appropriate. Similar to Caserta and de Vries (2005), we also model the tail of the distribution of the bidder's valuation as in the EVT setup.

Unfortunately, the message of the paper is somewhat bleak. We show that while for the distributions of valuations in the max-domain of attraction with positive extreme value index, the logarithm of the expected payoff can be estimated after application of a correction factor, the expected payoff cannot be estimated consistently. A somewhat similar result is obtained for the negative case. Only for a subset of distributions in the max-domain of attraction with zero extreme value index does a consistent estimator exist with a certain speed of convergence under a suitable second order condition.

The paper is organized as follows. In Section 2, the record and 2-record model is revisited. Section 3 demonstrates the EVT approach with positive, negative and zero extreme value index. For the zero case, a subclass model of the domain is introduced. Section 4 concludes this paper.

2 The bidding activities in Internet auction

IAs have some features that differentiate these from the standard auction mechanisms, but are otherwise just the internet version of known auctions. The differences pertain to the bidding systems and the termination rules. The Internet facilitates the use of two bidding systems simultaneously. Most of the IA sites allow for *manual* and *proxy bidding*. Manual bidding is similar to the first price open ascending bid in an English auction, while the proxy bidding procedure captures the second price sealed bid mechanism studied by Vickrey (1962). Proxy bidding proceeds by providing the server of the IA with the maximum value a person would be willing to pay. The machine then takes over and keeps on overbidding on behalf of the proxy bidder as long as the other bids are below this maximum. Regarding the termination rules, there are also two alternatives. One type of IA ends after a pre-announced fixed lapse of time, while the other type has a variable auto-extended termination time. Typical examples are the eBay auctions and the Amazon auctions. The eBay auctions have a fixed ending time. The Amazon type auctions use the auto-extension termination rule.[3] At the beginning of the Amazon type auction, an initial ending time is announced. If no bidding takes place during the last ten minutes, the auction stops at the announced ending time. But if

---

[3]Although www.amazon.com has terminated their auction platform, since they used the feature of auto-extension termination rule, we still call auctions with such kind of setup the Amazon type auction. On the Yahoo! platform, this feature is still in use.

there are some bids in the last ten minutes, the ending time is automatically extended by another ten minutes. This rule is also applied to the new extension period. On the Yahoo! auction site the sellers can choose between these two termination rules.

The analysis of auctions can be divided into two classes. Either it is assumed that the bidders' valuations are independent from each other, or they are dependent. The former case is usually referred to as the independent private values paradigm (IPVP). Valuations are considered to be draws from some given distribution. This is the paradigm that we consider in this paper as well. Standard commodities are well modeled on the IPVP assumption. Rare items, collectibles and works of art are usually considered to be in the other class. At the extreme end of the other class is the common value case. Under the IPVP and the Amazon type termination rule, de Haan et al. (2008) argued that the active bidders come to the IA as a record and 2-record arrival process, while their valuations form the record and 2-record sequence of the valuations among all potential bidders.

To explain this model, let $i = 1, 2, \cdots, n$ denote the order in which the n potential bidders arrive at the auction site. IPVP assumes that the valuation of all potential bidders are i.i.d. random variables $X_1, X_2, \cdots, X_n$ with distribution function $F(x)$. Define the rank sequence $\{R_i\}_{i=1}^n$ as

$$R_i := \sum_{k=1}^i 1_{\{X_k \geq X_i\}}. \tag{1}$$

Intuitively, R_i is the rank of the valuation of the i−th potential bidder among the valuations of all the potential bidders who checked the auction earlier (before i). The valuation X_i is called a *record* if $R_i = 1$. Similarly, for $k = 2, 3, \cdots$, it is a *k-record* if $R_i = k$, see Resnick (1987). Denote the indices of the records and 2-records as $\{J(j)\}_{j=1}^m$. This index sequence is given by

$$J(1) = 1, \ J(2) = 2 \tag{2}$$

$$J(j+1) = \min \{i > J(j) : R_i \leq 2\}, \ j = 2, 3, \cdots, m-1, \tag{3}$$

where m is the number such that $R_i > 2$ for all $i > J(m)$.

With the maintained hypothesis that *"each active (manual) bidder immediately returns to the IA and increases his bid as soon as he is overbid and his valuation is above the prevailing price."*,[4] the active bidders must have the indices $\{J(j)\}_{j=1}^m$ in the potential bidders sequence. So m is the number of active bidders. Then, the active bidders' valuations are obviously the record and 2-record sequence $\{X_{J(j)}\}_{j=1}^m$. Actually, the first $m-1$ active bidders' valuations can be observed as their last bids. The winner's valuation $X_{J(m)}$ is obviously unobservable, just as in the English auction. For the Amazon type auction, since there is no motivation for bidders to postpone their bids for strategic reasons, it can be assumed that the bids reflect the first $m-1$ records

---

[4]Note that this assumption is automatically satisfied when there are only proxy bidders present.

and 2-records. de Haan et al. (2008) tested this model by employing Yahoo! IA data. A similar record breaking model is considered in Bradlow and Park (2007). As an alternative, Shmueli et al. (2007) proposed an empirical approach on modeling the bidder's arrivals.

3 EVT approaches

Our purpose is to compare the observed payoff and its expectation for a specific IA. With n potential bidders, and m active bidders, there are two ways to represent the payoff following the record and 2-record model. One way is to consider $M_{n-1:n}$ as the second largest order statistics of $X_1, X_2, \cdots, X_n$. The other way is to view the payoff as the $X_{J(m-1)}$, where $\{J(j)\}_{j=1}^m$ is the record and 2-record index sequence as defined in the previous section.

Since the payoff is determined by the largest order statistics, it is reasonable to make assumptions only on the right tail of the valuation distribution $F(x)$. Caserta and de Vries (2005) suggested to use the EVT approach, and assumed that the distribution of the bidder's valuation belongs to the max-domain of attraction of an extreme value distribution. This setup is as follows.

Suppose the bidder's valuations are i.i.d random variables $X_1, X_2, \cdots,$ $X_n, \cdots$ with common distribution function F. Denote $M_n = \max\{X_1, \cdots, X_n\}$. We say that F belongs to the max-domain of attraction, if there exist a non-degenerate distribution function G, a positive sequence $\{a_n\}_{n=1}^\infty$ and a real sequence $\{b_n\}_{n=1}^\infty$, such that

$$\lim_{n \to \infty} P\left\{\frac{M_n - b_n}{a_n} \le x\right\} = G(x)$$

for all continuity points of G. Denote this domain of attraction feature as $F \in \mathcal{D}(G)$.

The necessary and sufficient condition for a distribution function to belong to the max-domain of attraction is the extreme value condition, see e.g. de Haan (1984).

Proposition 3.1 *Let $U := \left(\frac{1}{1-F}\right)^{\leftarrow}$ be the generalized inverse function of $1/(1-F)$. Then $F \in \mathcal{D}(G)$ if and only if there exists a function $a(t) > 0$ such that*

$$\lim_{t \to \infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}, \tag{4}$$

for some $\gamma \in \mathbb{R}$ and all $x > 0$.

Here γ is called the *extreme value index*. Under the extreme value condition, the following proposition is proved in Caserta and de Vries (2005).

Proposition 3.2 *Let $X_1, \cdots, X_n, \cdots$ be i.i.d. sequence with common distribution function F belonging to domain of attraction, i.e. Eq. 4 holds for some $\gamma < 0$. Then, we have that*

$$\lim_{n \to \infty} \frac{EM_{n-1:n} - U(n)}{a(n)} = -\Gamma(2 - \gamma) \tag{5}$$

From Proposition 3.2, a possible estimator of $EM_{n-1:n}$ is $\hat{U}(n) - \hat{a}(n)$ $\Gamma(2 - \hat{\gamma})$, where $\hat{U}(n)$, $\hat{a}(n)$,and $\hat{\gamma}$ are proper estimators for the location, scale and shape parameter in the EVT model. The literature offers several alternative estimations for these parameters. We note here that Eq. 5 holds for $0 \leq \gamma < 2$ as well.

Since in our model, only the record and 2-record sequence is observed, it is necessary to have proper estimators based on only those observations. When γ is positive, Berred (1992) derived an estimator for γ based on the record sequence, which can be generalized to our case of the record and 2-record sequence.

The main difficulty in this approach is that the number of potential bidders n is in fact unknown. This lack of information inhibits the estimation of $U(n)$ and $a(n)$. In Caserta and de Vries (2005) n is assumed to be equal to the number of bids. Thus multiple bids from the same bidder are considered as coming from different potential bidders. Although this estimate of n is larger than the observed m, it is a rather inaccurate estimation. In de Haan et al. (2008), it is shown that $m \sim 2 \log n$ as $n \to \infty$. All in all, to consider the expected payoff as the expectation of the second largest valuation seems to be an approach of limited value due to the unknown number of potential bidders.

That leaves us with the second representation to model the payoff based on the record theory. We first study the record and 2-record sequence via its point process representation. Let $\{L_k(n)\}_{n=1}^{\infty}$ be the indices of the $k-$record sequence, that is

$$L_k(1) = 1, \ L_k(n+1) = \min\left\{j > L_k(n) : R_j = k\right\}, \ n = 1, 2, \cdots.$$

Define the $k-$record point process N_k by

$$N_k := \sum_{n=1}^{\infty} \varepsilon_{X_{L_k(n)}},$$

for $k = 1, 2, \cdots$. Then according to Proposition 4.30 in Resnick (1987), the point processes $\{N_k\}_{k=1}^{\infty}$ are i.i.d. random elements.

It is clear that the combination of L_1 and L_2 sequences constitutes the record and 2-record sequence J. According to Proposition 4.1(ii) in Resnick (1987), N_1 and N_2 are homogeneous Poisson processes on $(0, +\infty)$. Then, the point process of the records and 2-records must be

$$N = N_1 + N_2$$

which is the sum of two independent homogeneous Poisson process on $(0, +\infty)$. So N is a Poisson process with intensity measure 2μ, where μ is the Lebesgue measure on $(0, +\infty)$. In other words, we have the following Lemma.

Lemma 3.1 *Suppose $E_1, E_2, \cdots$ are i.i.d standard exponentially distributed random variables and $\{J_E(m)\}$ is the index sequence of the records and 2-records of $\{E_n\}$. Then*

$$\left\{E_{J_E(m)}\right\}_{m=2}^{\infty} \stackrel{d}{=} \{\Gamma_m\}_{m=2}^{\infty}$$

where $\Gamma_m = \sum_{i=1}^{m} E_i'$ is the partial sum of the sequence $\left\{E_i'\right\}_{i=1}^{\infty}$ which is an i.i.d sequence with exponential distribution and mean $1/2$.

By defining $Q := (-\log(1 - F))^{\leftarrow}$, the i.i.d sequence $\{X_n\}$ can be represented as $\{Q(E_n)\}$, where $E_1, E_2, \cdots$ are i.i.d standard exponentially distributed random variables. Hence, a direct implication of Lemma 3.1 is as follows.

Corollary 3.1 *The record and 2-record sequence can be represented as*

$$\left\{X_{J(m)}\right\}_{m=2}^{\infty} \stackrel{d}{=} \{Q(\Gamma_m)\}_{m=2}^{\infty}$$

where $\Gamma_m = \sum_{i=1}^{m} E_i'$ is the partial sum of the sequence $\left\{E_i'\right\}_{i=1}^{\infty}$ which is an i.i.d sequence with exponential distribution and mean $1/2$.

By definition, the Q and U functions are connected by $Q(t) = U(e^t)$. Therefore, the extreme value condition in Eq. 4 can be rewritten in terms of the Q function as

$$\lim_{t \to \infty} \frac{Q(t + x) - Q(t)}{a(e^t)} = \frac{e^{\gamma x} - 1}{\gamma}, \tag{6}$$

where γ is the extreme value index. We discuss in the three separate subsections the cases $\gamma > 0$, $\gamma < 0$ and $\gamma = 0$.

3.1 Positive case: $\gamma > 0$

When Eq. 4 holds with $\gamma > 0$, we have that $\lim_{t\to\infty} U(t) = \infty$ and U is a regularly varying function at infinity, i.e.

$$\lim_{t \to \infty} \frac{U(tx)}{U(t)} = x^{\gamma}.$$

According to Proposition B.1.9 in de Haan and Ferreira (2006), $\log U(t)/\log t \to \gamma$. Hence for any $\delta > 0$, there exists $t(\delta) > 0$ such that for any $t > t(\delta)$,

$$t^{\gamma - \delta} < U(t) < t^{\gamma + \delta}.$$

Correspondingly, for all $t > t_1(\delta) := \log t(\delta)$, we have that

$$e^{(\gamma - \delta)t} < Q(t) < e^{(\gamma + \delta)t}. \tag{7}$$

Therefore, for $\delta < \gamma$, we have that

$$Q(\Gamma_m) > e^{(\gamma-\delta)\Gamma_m} 1_{\Gamma_m > t_1(\delta)} + Q(\Gamma_m) 1_{\Gamma_m \leq t_1(\delta)}$$

$$> e^{(\gamma-\delta)\Gamma_m} (1 - 1_{\Gamma_m \leq t_1(\delta)})$$

$$> e^{(\gamma-\delta)\Gamma_m} - e^{(\gamma-\delta)t_1(\delta)} \tag{8}$$

and

$$Q(\Gamma_m) < e^{(\gamma+\delta)\Gamma_m} 1_{\Gamma_m > t_1(\delta)} + Q(\Gamma_m) 1_{\Gamma_m \leq t_1(\delta)}$$

$$< e^{(\gamma+\delta)\Gamma_m} + Q(t_1(\delta)). \tag{9}$$

The expectation of $e^{\lambda \Gamma_m}$ for $\lambda \in \mathbb{R}$ is calculated in the following lemma.

Lemma 3.2 *For $\lambda \geq 2$ and any fixed integer m, $Ee^{\lambda \Gamma_m} = +\infty$. For $\lambda < 2$ and any fixed integer m*

$$Ee^{\lambda \Gamma_m} = \left(\frac{2}{2-\lambda} \right)^m.$$

When $\gamma > 2$, there exists a $\delta > 0$ such that $\gamma - \delta > 2$. By taking expectations at the two sides of Eq. 8, we get that

$$EQ(\Gamma_m) \geq Ee^{(\gamma-\delta)\Gamma_m} - e^{(\gamma-\delta)t_1(\delta)} = +\infty.$$

Therefore, we conclude that if $\gamma > 2$, for any finite level m, $EQ(\Gamma_m)$ does not exist. In other words, the current observed price necessarily underestimates the expected payoff. However, $\gamma > 2$ is not very realistic for most items offered on the IA platforms. For most, if not all, items the payoff does have an expectation. In case of a finite expected payoff, i.e. $\gamma < 2$, we have following theorem.

Theorem 3.1 *Suppose Eq. 6 holds for $0 < \gamma < 2$. Then $EQ(\Gamma_m)$ is finite for any fixed m. As $m \to \infty$, we have the following two limit relations*

$$\frac{EQ(\Gamma_m)}{Q(\Gamma_m)} \xrightarrow{P} +\infty, \tag{10}$$

and

$$\frac{\log EQ(\Gamma_m)}{\log Q(\Gamma_m)} \xrightarrow{P} c, \tag{11}$$

where $c := \frac{\log\left(\frac{2}{2-\gamma}\right)}{\gamma/2}$ is a constant larger than 1.

Proof of Theorem 3.1 Since Q is a monotone function, we only need to prove that $EQ(\Gamma_m)$ is finite for large m. Choose $\delta < \min(\gamma, 2 - \gamma)$. By taking expectations at the two sides of Eq. 9, we get that

$$EQ(\Gamma_m) \leq Ee^{(\gamma+\delta)\Gamma_m} + Q(t_1(\delta)) < +\infty.$$

Hence $EQ(\Gamma_m)$ is finite. Similarly, we have the lower bound of $EQ(\Gamma_m)$ as

$$EQ(\Gamma_m) \geq Ee^{(\gamma-\delta)\Gamma_m} - e^{(\gamma-\delta)t_1(\delta)} = \left(\frac{2}{2-\gamma+\delta}\right)^m - e^{(\gamma-\delta)t_1(\delta)}.$$

Together with Eq. 9, we find that

$$\frac{EQ(\Gamma_m)}{Q(\Gamma_m)} \geq \frac{\left(\frac{2}{2-\gamma+\delta}\right)^m - e^{(\gamma-\delta)t_1(\delta)}}{e^{(\gamma+\delta)\Gamma_m} + Q(t_1(\delta))} = \frac{1 - \left(\frac{2}{2-\gamma+\delta}\right)^{-m} e^{(\gamma-\delta)t_1(\delta)}}{e^{(\gamma+\delta)\Gamma_m - m\log\left(\frac{2}{2-\gamma+\delta}\right)} + \left(\frac{2}{2-\gamma+\delta}\right)^{-m} Q(t_1(\delta))}. \tag{12}$$

Denote $c(\delta) := 2\frac{\log\left(\frac{2}{2-(\gamma-\delta)}\right)}{\gamma+\delta}$. Since $\frac{2}{2-(\gamma-\delta)} > 1$, Eq. 12 is continued as

$$\frac{EQ(\Gamma_m)}{Q(\Gamma_m)} \geq \frac{1 - o(1)}{e^{(\gamma+\delta)(\Gamma_m - mc(\delta)/2)} + o(1)}. \tag{13}$$

Note that as $\delta \to 0, c(\delta) \to c$. From the inequality that $\log\frac{1}{1-x} > x$ for all $x < 1$, we get that $c > 1$. Hence, we can choose δ small enough such that $c(\delta) > 1$. From central limit theorem, we have that $\frac{\Gamma_m - m/2}{\sqrt{m}/2}$ is asymptotically standard normally distributed. Thus, for any $c(\delta) > 1$, $\Gamma_m - mc(\delta)/2 \xrightarrow{P} -\infty$ as $m \to \infty$. Therefore, as $m \to \infty$, the right side of Eq. 13 goes to $+\infty$ which completes the proof of Eq. 10.

From the boundaries of $EQ(\Gamma_m)$, we have that

$$\left(\frac{2}{2-\gamma+\delta}\right)^m - e^{(\gamma-\delta)t_1(\delta)} \leq EQ(\Gamma_m) \leq \left(\frac{2}{2-\gamma-\delta}\right)^m + Q(t_1(\delta)).$$

Hence, by taking logarithms and asking $m \to \infty$, we get

$$\log\left(\frac{2}{2-\gamma+\delta}\right) \leq \liminf_{m\to\infty} \frac{\log EQ(\Gamma_m)}{m}$$

$$\leq \limsup_{m\to\infty} \frac{\log EQ(\Gamma_m)}{m} \leq \log\left(\frac{2}{2-\gamma-\delta}\right).$$

By taking $\delta \to 0$, it follows that

$$\lim_{m\to\infty} \frac{\log EQ(\Gamma_m)}{m} = \log\left(\frac{2}{2-\gamma}\right). \tag{14}$$

Since $\lim_{t\to\infty} \log Q(t)/t \to \gamma$ and $\Gamma_m \xrightarrow{P} \infty$ as $m \to \infty$, we get that as $m \to \infty$

$$\log Q(\Gamma_m)/\Gamma_m \xrightarrow{P} \gamma.$$

From the Law of Large Numbers, we have that $\Gamma_m/(m/2) \xrightarrow{P} 1$ as $m \to \infty$. Thus,

$$\log Q(\Gamma_m)/m \xrightarrow{P} \gamma/2.$$

Together with Eq. 14, this complete the proof of Eq. 11. $\qquad\square$

Theorem 3.1 implies that, for $0 < \gamma < 2$, although the expected payoff is bounded, the observed payoff always underestimates its expectation if there are numerous active bidders. The following remark gives the essential reason for the underestimation.

Remark 3.1 Under the EVT model with $0 < \gamma < 2$, the comparison between $EQ(\Gamma_m)$ and $Q(\Gamma_m)$ is essentially a comparison between $Ee^{\gamma\Gamma_m}$ and $e^{\gamma\Gamma_m}$. From Jensen's inequality, we have that

$$Ee^{\gamma\Gamma_m} \geq e^{\gamma E\Gamma_m}.$$

From the Law of Large Numbers, $\Gamma_m \sim E\Gamma_m$ as $m \to \infty$. Thus, we intuitively see why $Q(\Gamma_m)$ underestimates $EQ(\Gamma_m)$.

A question is whether it is possible to correct the underestimation. Theorem 3.1 shows that the logarithm of the expected payoff can be approximated by the logarithm of the observed payoff multiplied with an adjustment factor c that is always higher than 1. Notice that c is a function of the extreme value index γ which can be consistently estimated as a function of the observed record sequence, see Berred (1992). By estimating c, a consistent estimator for the logarithm of the expected payoff can be constructed. However, a consistent estimate at the log-level does not provide a consistent estimator for the expected payoff itself because both of the expected payoff and the observed payoff go to infinity as the number of active bidders m go to infinity. The situation is similar to the $2 \log n$ rule in de Haan et al. (2008). Given the number of potential bidders n, the number of active bidder m is consistently estimated as $2 \log n$. However, the $2 \log n$ rule does not provide a consistent estimator for the number of potential bidders given the number of active bidders. Therefore, within the framework of Theorem 3.1, it is not possible to have a consistent estimator of the expected payoff based on the observed record and 2-record sequence.

3.2 Negative case: $\gamma < 0$

In case $\gamma < 0$, the distribution function of the bidders' valuations F has a right endpoint, i.e. $Q(\infty) := \lim_{x\to+\infty} Q(x) < \infty$. Hence, the bidders' valuations are never above $Q(\infty)$.[5] In such a case, the expected payoff is always finite. Caserta and de Vries (2005) argued that this is a realistic model for most items sold through IA. For example, the new price is often a realistic upper bound of a second-hand consumer item sold through IA.

Since $\Gamma_m \overset{P}{\to} +\infty$ as $m \to \infty$, we get that $Q(\Gamma_m) \overset{P}{\to} Q(\infty)$. The following theorem studies the asymptotic difference between $Q(\Gamma_m)$ and $EQ(\Gamma_m)$.

---

[5] We remark that the uniform distribution, which is a commonly used distribution in auction theory, belongs to this case with $\gamma = -1$.

Theorem 3.2 *Suppose Eq. 6 holds for $\gamma < 0$. Then $EQ(\Gamma_m) \to Q(\infty)$ as $m \to \infty$. Thus,*

$$\frac{EQ(\Gamma_m)}{Q(\Gamma_m)} \xrightarrow{P} 1.$$

Furthermore, we have the following two limit relations as $m \to \infty$,

$$\frac{Q(\infty) - EQ(\Gamma_m)}{Q(\infty) - Q(\Gamma_m)} \xrightarrow{P} +\infty, \tag{15}$$

and

$$\frac{\log(Q(\infty) - EQ(\Gamma_m))}{\log(Q(\infty) - Q(\Gamma_m))} \xrightarrow{P} c, \tag{16}$$

where c is defined as in Theorem 3.1. Notice that for negative γ we have $c < 1$.

We start by proving the following useful lemma.

Lemma 3.3 *For any fixed constants $T > 0$, $\varepsilon > 0$,*

$$\lim_{m \to \infty} e^{\varepsilon m} P(\Gamma_m \leq T) = 0$$

Proof of Lemma 3.3 Notice that $2\Gamma_m$ follows a Gamma distribution with shape parameter m, i.e. the density function of $2\Gamma_m$ is $f(x) := \frac{x^{m-1}e^{-x}}{\Gamma(m)} 1_{x>0}$. We have that

$$
\begin{aligned}
0 < e^{\varepsilon m} P(\Gamma_m \leq T) &= e^{\varepsilon m} \int_0^{2T} \frac{x^{m-1}e^{-x}}{\Gamma(m)} dx \\
&= \int_0^{2T} \frac{(e^{\varepsilon}x)^{m-1}e^{-x}}{\Gamma(m)} d(e^{\varepsilon}x) \\
&= \int_0^{2T} e^{(e^{\varepsilon}-1)x} \frac{(e^{\varepsilon}x)^{m-1}e^{-(e^{\varepsilon}x)}}{\Gamma(m)} d(e^{\varepsilon}x) \\
&\leq e^{(e^{\varepsilon}-1)2T} \int_0^{2Te^{\varepsilon}} \frac{x^{m-1}e^{-x}}{\Gamma(m)} dx \\
&= e^{(e^{\varepsilon}-1)2T} P\left(\Gamma_m' \leq 2Te^{\varepsilon}\right)
\end{aligned}
$$

where Γ_m' is a Gamma distributed random variable with the same density function f. Because $\Gamma_m' \to +\infty$ as $m \to \infty$, the lemma is proved. □

Proof of Theorem 3.2 When $\gamma < 0$, $1/(Q(\infty) - Q(t))$ is a regularly varying function at $+\infty$ with index $-\gamma$. Similar to the inequality (7) in the positive case, the following inequality holds. For any $\delta > 0$, there exists $t_2(\delta)$ such that for any $t > t_2(\delta)$

$$e^{(\gamma-\delta)t} < Q(\infty) - Q(t) < e^{(\gamma+\delta)t}. \tag{17}$$

Therefore, we have that

$$Q(\infty) - Q(\Gamma_m) > e^{(\gamma-\delta)\Gamma_m} 1_{\Gamma_m > t_2(\delta)} + (Q(\infty) - Q(\Gamma_m)) 1_{\Gamma_m \leq t_1(\delta)}$$
$$> e^{(\gamma-\delta)\Gamma_m} \left(1 - 1_{\Gamma_m \leq t_2(\delta)}\right)$$
$$> e^{(\gamma-\delta)\Gamma_m} - 1_{\Gamma_m \leq t_2(\delta)}. \qquad (18)$$

and

$$Q(\infty) - Q(\Gamma_m) < e^{(\gamma+\delta)\Gamma_m} 1_{\Gamma_m > t_2(\delta)} + (Q(\infty) - Q(\Gamma_m)) 1_{\Gamma_m \leq t_2(\delta)}$$
$$< e^{(\gamma+\delta)\Gamma_m} + Q(\infty) 1_{\Gamma_m \leq t_2(\delta)}. \qquad (19)$$

By taking expectations at the two sides of Eq. 19, we get that

$$0 < Q(\infty) - EQ(\Gamma_m) \leq \left(\frac{2}{2 - \gamma - \delta}\right)^m + Q(\infty) P(\Gamma_m \leq t_2(\delta)).$$

By taking $\delta < -\gamma$, we get that $\frac{2}{2-\gamma-\delta} < 1$. Since $P(\Gamma_m \leq t_2(\delta)) \to 0$ as $m \to \infty$, we get that $\lim_{m\to\infty} EQ(\Gamma_m) = Q(\infty)$.

We turn to compare $Q(\infty) - EQ(\Gamma_m)$ with $Q(\infty) - Q(\Gamma_m)$. By taking expectations on the two sides of Eq. 18, we get that

$$Q(\infty) - EQ(\Gamma_m) \geq \left(\frac{2}{2 - \gamma + \delta}\right)^m - P(\Gamma_m \leq t_2(\delta)).$$

Notice that $\Gamma_m \xrightarrow{P} +\infty$ as $m \to \infty$. The inequality (19) implies that eventually

$$Q(\infty) - Q(\Gamma_m) < e^{(\gamma+\delta)\Gamma_m}.$$

Hence

$$\frac{Q(\infty) - EQ(\Gamma_m)}{Q(\infty) - Q(\Gamma_m)} \geq \frac{\left(\frac{2}{2-\gamma+\delta}\right)^m - P(\Gamma_m \leq t_2(\delta))}{e^{(\gamma+\delta)\Gamma_m}}. \qquad (20)$$

Lemma 3.3 shows that $P(\Gamma_m \leq t_2(\delta))$ goes to 0 at a higher speed than any exponential speed. Thus, Eq. 20 is continued as

$$\frac{Q(\infty) - EQ(\Gamma_m)}{Q(\infty) - Q(\Gamma_m)} \geq \frac{1 - o(1)}{e^{(\gamma+\delta)(\Gamma_m - d(\delta)m/2)}}, \qquad (21)$$

where $d = 2\frac{\log\left(\frac{2}{2-\gamma+\delta}\right)}{\gamma+\delta}$. Notice that as $\delta \to 0$, $d(\delta) \to c$ and $c < 1$ holds for $\gamma < 0$. Hence, we can choose δ small enough such that $d(\delta) < 1$. Then, $\Gamma_m - d(\delta)m/2 \xrightarrow{P} +\infty$ as $m \to \infty$. Therefore, as $m \to \infty$, the right side of Eq. 21 goes to $+\infty$ which completes the proof of the theorem. The proof of Eq. 16 is similar to that of Eq. 11. $\qquad \square$

Theorem 3.2 studies the case $\gamma < 0$ and tells a story just opposite to the positive case. For $\gamma < 0$, the observed payoff might be considered as a consistent estimator of its expectation because both the observed and expected payoff converge to the right endpoint of the bidder's valuation. Nevertheless,

the distance between the observed payoff and the right endpoint is eventually smaller than the distance between the expected payoff and the right endpoint. Hence, if there are sufficiently many active bidders, the observed payoff always overestimates its expectation. A consistent estimator on the log-level of the difference is given by multiplication with an adjustment factor c that is always lower than 1. However, similar to the positive case, a consistent estimate at the log-level would not provide a consistent estimator of the difference itself. Therefore, it is not possible to correct the overestimation.

To sum up, Theorems 3.1 and 3.2 show that if the extreme value index of bidder's valuation distribution is not 0, the observed payoff is never a satisfactory estimator for the expected payoff.

3.3 Zero case: $\gamma = 0$

In the previous two Subsections, we found that the observed payoff underestimates or overestimates its expectation when γ is positive or negative respectively. The remaining case is $\gamma = 0$, i.e. when F belongs to the Gumbel domain. In this section, we show that at least for a subclass of the Gumbel domain, the observed payoff is a reasonable estimator for its expectation.

3.3.1 Model specification

As we discussed before, in order to avoid the problem of the unknown number of potential bidders, the model should be based on the Q function and the payoff should be taken as the last observation in the record and 2-record sequence.

We introduce a refinement and assume that the Q function itself is regularly varying or generalized regularly varying. Furthermore, in order to study the asymptotic properties, we assume that a second-order condition holds.

We start from the regularly varying model. Suppose Q function itself is regularly varying with index $\lambda > 0$. We also assume that it is second-order regularly varying with second-order index $\rho \leq 0$, i.e.

$$\frac{\frac{Q(tx)}{Q(t)} - x^\lambda}{A(t)} \to H(x) := x^\lambda \frac{x^\rho - 1}{\rho} \tag{22}$$

as $t \to \infty$, for some suitable function $A(t) \in RV_\rho$ and all $x > 0$. We call this *Regularly Varying Q-function (RVQ)* model.

Remark 3.2 The RVQ model with $\rho < -1$ is a subclass of the Gumbel domain.

Proof of Remark 3.2 From Eq. 22, we have the following inequality (See de Haan and Ferreira (2006, Appendix B)). Given any $\varepsilon > 0$, there is a $t_0(\varepsilon)$ such that for all $tx > t_0(\varepsilon)$,

$$\left| \frac{\frac{Q(tx)}{Q(t)} - x^\lambda}{A(t)} - H(x) \right| \leq \varepsilon x^{\lambda + \rho + \varepsilon}. \tag{23}$$

Therefore,

$$\left| \frac{\frac{Q(t+x)}{Q(t)} - \left(\frac{t+x}{t}\right)^\lambda}{A(t)} - H\left(\frac{t+x}{t}\right) \right| \le \varepsilon \left(\frac{t+x}{t}\right)^{\lambda+\rho+\varepsilon},$$

for all $t > t_0(\varepsilon)$ and positive x. When $\rho < -1$ and $A(t) \in RV_\rho$, we get $tA(t) \to 0$ as $t \to \infty$. It leads to the fact that

$$t\left\{ \frac{Q(t+x) - Q(t)}{Q(t)} - ((1 + x/t)^\lambda - 1) \right\} \to 0.$$

Since

$$t((1 + x/t)^\lambda - 1) \to \lambda x,$$

we finally get that

$$\frac{Q(t+x) - Q(t)}{\lambda Q(t)/\sqrt{t}} \to x,$$

i.e. the corresponding U function satisfies condition Eq. 4 with $\gamma = 0$. □

Thus, the RVQ model is only a special case in the Gumbel domain ($\gamma = 0$), and therefore is more narrow. Fortunately, quite some well-known parametric distributions belong to this model. For example, both the normal distribution and the exponential distribution satisfy this model. (For the normal distribution, $\lambda = 1/2$, $\rho = -\infty$. For the exponential distribution, $\lambda = 1$, $\rho = -\infty$.)

In the RVQ model, the original distribution function can not have a finite right endpoint. In order to include distributions with finite right endpoint, we extend the model as follows. Suppose the Q function is second-order generalized regularly varying with first-order index $\lambda \in \mathbb{R}$ and second-order index $\rho \le 0$, i.e.

$$\frac{\frac{Q(tx)-Q(t)}{a(t)} - \frac{x^\lambda - 1}{\lambda}}{A(t)} \to H(x) \qquad (24)$$

as $t \to \infty$, for all $x > 0$ and some suitable function $a(t) \in RV_\lambda$, $A(t) \in RV_\rho$ and $H(x)$. We call this *Generalized Regularly Varying Q-function (GRVQ)* model.

In the GRVQ model, if $\lambda > 0$, it can be simplified to the RVQ model. When λ is negative, $Q(\infty) < \infty$. Hence, in this case, the original distribution function F must have a finite right endpoint.[6] We note though that a random variable K with distribution function

$$1 - e^{-((a-x)/b)^{1/\lambda}}$$

---

for all $x \leq a$, where $b > 0$, $\lambda < 0$ satisfies the requirements of the GRVQ model (λ is the regularly varying index in Eq. 24 and $\rho = -\infty$). Notice that $-K$ follows the inverse Weibull distribution.

3.3.2 Statistical inference

We turn to study the expected payoff. The following lemma gives the asymptotic properties of the observed and expected payoff.

Lemma 3.4 *Suppose the RVQ model holds with $\rho < -1/2$, then*

$$\sqrt{m} \left(\frac{X_{J(m)}}{Q(m/2)} - 1 \right) \xrightarrow{d} \lambda W, \tag{25}$$

$$\sqrt{m} \left(\frac{EX_{J(m)}}{Q(m/2)} - \frac{\Gamma(\lambda + m)}{m^\lambda \Gamma(m)} \right) \to 0, \tag{26}$$

where W is a standard normal distributed random variable and Γ is the Gamma function.

Proof of Lemma 3.4 Applying Eq. 23, we get that, eventually,

$$\left| \frac{\frac{Q(\Gamma_m)}{Q(m/2)} - \left(\frac{\Gamma_m}{m/2} \right)^\lambda}{A(m/2)} - H \left(\frac{\Gamma_m}{m/2} \right) \right| \leq \varepsilon \left(\frac{\Gamma_m}{m/2} \right)^{\lambda + \rho \pm \varepsilon}. \tag{27}$$

The symbol $\pm$ means taking the suitable sign according to whether $\frac{\Gamma_m}{m/2}$ is higher or lower than 1. Since $\sqrt{m} A(m/2) \to 0$, $\frac{\Gamma_m}{m/2} \xrightarrow{P} 1$ and $\lim_{x \to 1} H(x) \pm \varepsilon x^{\lambda + \rho \pm \varepsilon}$ exists, we find that

$$\sqrt{m} \left(\frac{Q(\Gamma_m)}{Q(m/2)} - \left(\frac{\Gamma_m}{m/2} \right)^\lambda \right) \xrightarrow{P} 0$$

From Central Limit Theory, we have that

$$\sqrt{m} \left(\frac{\Gamma_m}{m/2} - 1 \right) \xrightarrow{d} W,$$

where W is a standard normal distributed random variable. According to Cramèr's delta method,

$$\sqrt{m} \left(\left(\frac{\Gamma_m}{m/2} \right)^\lambda - 1 \right) \xrightarrow{d} \lambda W.$$

Thus, Eq. 25 is a direct consequence.

Taking expectation of the two sides of inequality (27) on the set $\{\Gamma_m > t_0(\varepsilon)\}$, because the absolute value function is a convex function, we get that

$$\left| \frac{\frac{EQ(\Gamma_m)1_{\{\Gamma_m > t_0(\varepsilon)\}}}{Q(m/2)} - E\left(\frac{\Gamma_m}{m/2}\right)^\lambda 1_{\{\Gamma_m > t_0(\varepsilon)\}}}{A(m/2)} - EH\left(\frac{\Gamma_m}{m/2}\right)1_{\{\Gamma_m > t_0(\varepsilon)\}} \right|$$

$$\leq \varepsilon E\left(\frac{\Gamma_m}{m/2}\right)^{\lambda+\rho\pm\varepsilon} 1_{\{\Gamma_m > t_0(\varepsilon)\}}.$$

Similar to above discussion, we can conclude that

$$\sqrt{m}\left(\frac{EQ(\Gamma_m)1_{\{\Gamma_m > t_0(\varepsilon)\}}}{Q(m/2)} - E\left(\frac{\Gamma_m}{m/2}\right)^\lambda 1_{\{\Gamma_m > t_0(\varepsilon)\}} \right) \to 0. \tag{28}$$

Since $\sqrt{m}P(\Gamma_m \leq x) \to 0$ as $m \to \infty$ for all $x > 0$, we have that

$$\sqrt{m}\left(\frac{EQ(\Gamma_m)1_{\{\Gamma_m \leq t_0(\varepsilon)\}}}{Q(m/2)} \right) \leq \sqrt{m}\left(\frac{Q(t_0(\varepsilon))P(\{\Gamma_m \leq t_0(\varepsilon)\})}{Q(m/2)} \right) \to 0,$$

and

$$\sqrt{m}E\left(\frac{\Gamma_m}{m/2}\right)^\lambda 1_{\{\Gamma_m \leq t_0(\varepsilon)\}} \leq \sqrt{m}E\left(\frac{t_0(\varepsilon)}{m/2}\right)^\lambda P(\{\Gamma_m \leq t_0(\varepsilon)\}) \to 0.$$

It follows that Eq. 28 can be rewritten as

$$\sqrt{m}\left(\frac{EQ(\Gamma_m)}{Q(m/2)} - E\left(\frac{\Gamma_m}{m/2}\right)^\lambda \right) \to 0 \tag{29}$$

Since $2\Gamma_m$ follows the Gamma distribution with shape parameter m, the expectation of $(2\Gamma_m)^\lambda$ is $\frac{\Gamma(\lambda+m)}{\Gamma(\lambda)}$ (see e.g., Papoulis (1984), pp. 103–104). Then

$$E\left(\frac{\Gamma_m}{m/2}\right)^\lambda = \frac{\Gamma(\lambda+m)}{m^\lambda \Gamma(m)}.$$

Thus, Eq. 26 is proved. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Now, we prove that

$$\lim_{m\to\infty} \sqrt{m}\left(\frac{\Gamma(\lambda+m)}{m^\lambda \Gamma(m)} - 1 \right) = 0. \tag{30}$$

Since $\Gamma(x) = e^{-x}x^{x-1/2}\sqrt{2\pi}(1 + 1/12xo(1)))$ as $x \to \infty$, we have that

$$\frac{\Gamma(\lambda + m)}{m^\lambda \Gamma(m)} = e^{-\lambda}(1 + \lambda/m)^{m-1/2}(1 + \lambda/m)^\lambda \frac{1 + \frac{\lambda+m}{12}o(1)}{1 + \frac{m}{12}o(1)}$$

$$= e^{(m-1/2)\log(1+\lambda/m)-\lambda}\left(1 + \lambda^2/m + o(1/m)\right)(1 + o(1/m))$$

$$= e^{-\frac{\lambda-\lambda^2}{2m}+o(1/m)}\left(1 + \lambda^2/m + o(1/m)\right)(1 + o(1/m))$$

$$= 1 + O(1/m).$$

Thus Eq. 30 is proved.

We can estimate the expected payoff as follows. Suppose we have m active bidders in an IA. We observe the bidders' valuations $X_{J(1)}, \cdots, X_{J(m-1)}$ as their final bids, except for the winner. The payoff will be $X_{J(m-1)}$. We estimate the expected payoff $EX_{J(m-1)}$ by this observation. The asymptotic property of this estimator is given by the following theorem.

Theorem 3.3 *Suppose the RVQ model holds for $\rho < -\frac{1}{2}$. Then*

$$\sqrt{m-1}\left(\frac{X_{J(m-1)}}{EX_{J(m-1)}} - 1\right) \xrightarrow{d} \lambda W \tag{31}$$

as $m \to \infty$, where W is a standard normally distributed random variable.

Proof of Theorem 3.3 The theorem is proved by combining Eqs. 25, 26 and 30. □

Theorem 3.3 shows that for the RVQ model the observed payoff is an accurate estimator for the expected payoff and the corresponding asymptotic normality holds under a second order condition.

Starting from the GRVQ model, similar results can be obtained as in the previous subsection. Here we only present the conclusion, the proof is omitted. The proof for the GRVQ model is essentially the same as the proof for the RVQ model.

Theorem 3.4 *Suppose the GRVQ model holds with $\rho < -\frac{1}{2}$. Then as $m \to \infty$,*

$$\sqrt{m-1}\frac{X_{J(m-1)} - EX_{J(m-1)}}{a((m-1)/2)} \xrightarrow{d} W,$$

where W is a standard normally distributed random variable. In particular, when $\lambda > 0$, we have that Eq. 31 holds. When $\lambda < 0$, we have that $Q(\infty) < \infty$, and as $m \to \infty$,

$$\sqrt{m-1}\left(\frac{Q(\infty) - X_{J(m-1)}}{Q(\infty) - EX_{J(m-1)}} - 1\right) \xrightarrow{d} \lambda W.$$

4 Conclusion

Internet auctions are a hybrid of the standard second-price Vickrey auction and the first-price English auction, for which the expected payoff is equal to the expectation of the second highest valuation among all the potential bidders. The expected payoff acts as a benchmark for the reasonableness of the price that is paid for the purchased item. Since the number of potential bidders is not observable, the expected value is difficult to estimate accurately. We approached this problem by considering the bids as a record and 2-record sequence of the potential bidder's valuation. The observed payoff is thus one of the records and 2-records.

In this paper, we use the EVT models to model the tail distribution of the bidder's valuation and to study the expected payoff. We first argue that assuming that the extreme value index γ is higher than 2 is not a realistic model because in that case the expected payoff is unbounded. For $0 < \gamma < 2$, we show that the observed payoff underestimates the expected payoff. At the log-level, an adjusted estimator exists for the expected payoff based on the logarithm of the observed payoff. We show that this is not possible at the level of the expected payoff. Hence, the consistency is at stake. One may argue that $\gamma > 0$ is not a realistic setup for the distribution of the bidder's valuation, because such a distribution function has no finite right endpoint which may not reflect the reality of IA.

For $\gamma < 0$ the distribution function of the bidder's valuation has a finite endpoint, which may be a more realistic setup for IA. Both the expected and observed payoff converge to the right endpoint as the number of active bidders m goes to infinity. However, the distances to the endpoint converge to 0 at different speeds. The distance between the observed payoff and the right endpoint goes to 0 faster. Therefore, the observed payoff always overestimates the expected payoff.

For $\gamma = 0$, i.e. the distribution function of the bidder's valuation belongs to the Gumbel domain, the observed payoff can be a consistent estimator for the expected revenue. We introduced a subclass of the Gumbel domain as the model of bidder's valuation distribution. Within this subclass and under a second order condition, the observed payoff consistently converges to the expected payoff and the corresponding asymptotic normality holds.

All in all, in an IA the observed payoff is the final price of the deal, while the expected payoff is what the seller should get from holding such an IA. Our study shows that by assuming that the tail of the bidder's valuation distribution belongs to the domain of attraction of an extreme value distribution, the final price does not always reflect what the seller deserves.

Our analysis focused on the general EVT setup for the distribution of bidder's valuation. In case one is willing to make an explicit parametric assumption regarding the distribution of valuations, consistent estimation of the expected payoff may be possible outside the limited class that we could handle.

Further research on estimating the extreme value index from the observed record and 2-record sequence is also of interest and may help to identify the particular situation for a specific IA.

References

Bajari, P., Hortaçsu, A.: Winner's curse reserve prices and endogenous entry: empirical insights from eBay auction. RAND J. Econ. **34**, 329–355 (2003)

Berred, M.: On record values and the exponent of a distribution with regularly varying upper tail. J. Appl. Probab. **29**, 575–586 (1992)

Bradlow, E.T., Park, Y.H.: Bayesian estimation of bid sequences in Internet auctions using a generalized record breking model. Mark. Sci. **26**, 218–229 (2007)

Bulow, J.I., Klemperer, P.D.: Auctions vs. negotiations. Am. Econ. Rev. **86**, 180–194 (1996)

Caserta, S., de Vries, C.G.: Auctions with numerous bidders. Tinbergen Institute Discussion Paper, 05-031/2 (2005)

de Haan, L.: Slow variation and characterization of domains of attraction. In: de Oliveira, J.T. (ed.) Statistical Extremes and Application, pp. 31–38. Reidel, Dordrecht (1984)

de Haan, L., de Vries, C., Zhou, C.: The extent of Internet auction market. Tinbergen Institute Discussion Paper, 08-041/2 (2008)

de Haan, L., Ferreira, A.: Extreme Value Theory: An Introduction. Springer, New York (2006)

Krishna, V.: Auction Theory. Academic, London (2002)

Laffont, J., Ossard, H., Vuong, Q.: Econometrics of first price auctions. Econometrica **63**, 953–980 (1995)

McAfee, R.P., McMillan, J.: Auctions with a stochastic number of bidders. J. Econ. Theory **43**, 1–19 (1987)

Paarsch, H.J.: Deciding between the common and private value paradigms in empirical models of auctions. J. Econ. **51**, 191–251 (1992)

Papoulis, A.: Probability, Random Variables, and Stochastic Processes, 2nd edn. McGraw-Hill, New York (1984)

Resnick, S.I.: Extreme Values, Regular Variation, and Point Processes. Springer, New York (1987)

Shmueli, G., Russo, R.P., Jank, W.: The barista: a model for bid arrivals in online auctions. Ann. Appl. Statist. **1**, 412–441 (2007)

Song, U: Nonparametric estimation of an eBay auction model with an unknown number of bidders. Working paper. http://faculty.arts.ubc.ca/usong/ (2004)

Vickrey, W.: Auctions and bidding games. In: Recent Advances in Game Theory, pp. 15–27. The Princeton University Press, Princeton (1962)