Bayesian Learning at eBay?
Updating From Related Data and Empirical Evidence.

Katharina Sailer*

This Version: February 2006

PRELIMINARY - DO NOT CITE

Abstract

This paper introduces a model of Bayesian learning in second price auctions in which bidders belief to face a stationary distribution of competitors' bids over time. This reflects a situation which is typical for internet auctions like eBay. Learning in second price auctions is complicated by the fact that bidders cannot directly observe the statistic of interest, namely the highest bid of the competitors, but only the transaction price, which is the second highest bid. The paper shows that the latter can be used to update the beliefs about the location parameter of the distribution of the highest bid when relying on asymptotic distributions for the extremes. Panel estimations with data from eBay auctions show that learning provides an explanation for the observed bidding patterns.

1 Introduction

When looking at eBay data it strikes that bidders increase their bid with each trial in a new auction for the same object. Static auction models cannot explain this behavior. While the sequential auction literature (see e.g. Weber (2000)) can rationalize increasing individual bids,

*Katharina Sailer, Munich Graduate School in Economics and Center for Information and Network Economics, University of Munich, Akademiestr. 1/I, 80799 Muenchen, Germany. E-mail: ksailer@stat.uni-muenchen.de. I am grateful to Sven Rady, Joachim Winter, and Stefan Mittnik for their advice and encouragement. I would like to thank the Volkswagen Stiftung, the Nixdorf Stiftung, and the German Research Foundation for the generous financial support which made this research possible.

the assumptions which drive this behavior provide no good description for the eBay market. Sailer (2005) introduces a model of search which captures the eBay specific form of competition better. This paper shows that learning is a possible explanation for bidders' increasing bids in such an environment.

Why are standard sequential auction models not applicable to eBay? These models have a finite horizon, determined by the ex ante known number of products to be sold, and the number of bidders is always bigger than the available products. The reason why a bidder increases his bid over time is intimately related to this scarcity of supply: Each time he looses, his remaining winning chances decrease, and he, thus, bids more aggressively. In eBay's product categories for off-the-shelf products like computers, consumer electronics, or DVDs supply is rarely scarce in this form. While at a certain point in time the bidder sees only a limited number of open auctions, he knows, new ones will open soon. Likewise, new bidders permanently enter while others leave. The framework in Sailer (2005) allows for an infinite horizon and entry of new bidders. Instead of being restricted by a limited supply, bidders here refrain from bidding too low on account of positive bidding costs. Under these circumstances it is optimal for a bidder to bid his valuation less his continuation value - the same as in a sequential auction model. Given that the expected future relation between supply and demand remains stable the continuation value though does not change over time and the bidder places the same 'reservation bid' in each auction.

Increasing bids can arise in this environment when bidders engage in learning. Bidders probably are aware that products can be acquired at eBay cheaper than elsewhere. By exactly how much cheaper might be less clear to many of them when bidding for the first time. The price a bidder pays in a second price auction like eBay is determined by the highest bid of the competitors. A key simplifying assumption in Sailer (2005) is that bidders belief to play against an invariant distribution of competitors' bids over time.[1] If this is so, the bidder can use past observed bids to update his prior beliefs about this distribution. After each auction the bidder then re-assesses his future expected return which changes his continuation value which in turn affects the bidding strategy in the next auction. How exactly the uncertainty resolves over the

[1]Stationarity makes sense in a mature market combined with an environment where entry of new bidders, individual participation decisions, and stochastic components in valuations introduce so much noise that learning about the characteristics of any specific competitor after observing him in one auctions provides minor payoff.

course of a bidders' participation in the eBay game and in which way bidders can incorporate new information into their bidding strategy will be explored in the following.

Burdett and Vishwanath (1988) show how to introduce learning in a search model. In their model the job seeker in the labor market is uncertain about (the location parameter of) the job offer distribution. With each offer he receives, he updates his prior in a Bayesian way; consecutive decisions are based on the posterior distribution. The corresponding distribution in a second price auction is the distribution of the highest bid of the competitors which determines a bidder's winning odds and the price he pays. The main technical challenge is that bidders in these auctions normally do not observe the highest bid of their competitors and thus cannot use this information to update their beliefs. They only observe the transaction price, which is equal to the second highest bid of the competitors (+ 1 increment), and maybe some or all lower ordered bids. I will show in the following that a function of the transaction price provides a sufficient statistic for updating the distribution of the highest bid when relying on asymptotic distributions of order statistics. Moreover, starting from a prior knowledge of the distribution of the highest bid of the competitors, the bidder can update his beliefs each time he participates in an auction in a simple way, that is, using conjugate priors.

The statistical literature on asymptotic distributions of order statistics, also referred to as extreme value distributions, dates back to the beginnings of the last century. A seminal work is Gumbel (1958). It has been applied in so diverse fields as survival analysis in biology, reliability studies in engineering, or to measure any kind of extremal event, such as floods or wind speeds. The only application to economics, to my knowledge, is in finance where it is used to model the tails of the distribution of stock returns. The paper shows how to make use of the framework in a dynamic auction environment. In standard auctions the number of bidders is normally not big enough to justify the application of asymptotic distributions. This is different for internet auctions such as eBay; here the potential number of bidders, that is, those that consider bidding in any given auction, is for many product categories rather large.

As opposed to eBay.com, at eBay.de auctions cannot be searched for anymore as soon as they are closed. By this policy eBay makes it more difficult for a bidder to inform himself about his competitors' bids before participating in an auction. It can therefore be expected that part of the learning process a bidder undergoes can actually be observed in the bidding data. I will use a data set of all auctions and the corresponding bids for a specific palm pilot, which was collected from eBay.de during a seven month period, to see whether bidders combine their prior knowledge with the new information, obtained in those auctions where they participated in, in the way predicted by the model.

The model builds on the classic theoretical decision making literature which gives clear advice on how people should behave under uncertainty. Researchers have questioned whether people actually follow this advice. There is some evidence in the experimental literature that people are applying Bayesian rules to resolve uncertainty (see e.g. El-Gamal and Grether (1995)). The paper adds to this literature, using field data, by showing that bidders at eBay incorporate new information into their strategies in a Bayesian way.

The next section states a model of updating in second price auctions where a large number of potential bidders are present and bidders are unable to follow a specific bidder or a specific sample of bidders over time. This model reflects a situation which is typical for internet auctions. Section 3 discusses the impact of learning on the optimal bidding strategies. The application to eBay and estimation results are given in section 4. The last section concludes.

2 Benchmark Model

Setup. Assume an infinite number of homogenous products which are offered to n potential buyers in sequence, one in each period t. Whenever one buyer leaves the market for good, a new one enters. The products are sold via independent Vickrey auction. Bidder i is interested in one product only which he values at $v_{it} = v_i + \epsilon_{it}$. v_{it} is private information. While v_i remains constant over time, ϵ_{it} is drawn anew before the start of each auction from a common density function f_ϵ with mean zero and variance σ_ϵ.[2] The bidder can bid in as many auction as he wishes. Bidding though comes at a cost c which is common to all bidders. In each period the bidder has to decide about participation, $\delta_{it} \in \{0, 1\}$, and about a bidding policy, b_{it}. Policies are chosen such as to maximize the intertemporal utility given by the sum of the expected period returns.

[2]This form for the valuations is chosen in view of the empirical part. ϵ_{it} there reflects the time-varying part in the individual valuations which cannot be attributed to observables such as product characteristics.

Beliefs. To simplify the analysis and to focus the view on the learning aspect, I assume that the bidder believes to face the same distribution of competitors' bids in each auction. Sailer (2005) provides a detailed discussion of this assumption. When deriving the bids distribution from the underlying distribution of valuations it basically means two things: First, it excludes that a bidder learns about the personal characteristics - here the time-constant part v_j - of any specific competitor j by past observations. Second, the v_j's and ϵ_{jt}'s of all of the $n-1$ competitors are believed to come from time-invariant distributions, that is entry and exit happen in a way so that these distributions remain stationary. When concentrating on Markov equilibria, the only variable which affects bidding strategies is then the private valuation of a bidder, or more formally: $b_{jt} = b(v_{jt})$.[3] The bid of each potential competitor is therefore distributed according to some density f_b with corresponding distribution F_b which is composed of the densities of v and ϵ and is consequently believed not to change across auctions.

What really matters for the bidder is the highest bid of the actual participants in an auction since this determines his winning odds and the price he pays. The optimal Markovian participation decision of each potential competitor is given by $\delta_{jt} = \delta(v_{jt}) = \delta(b^{-1}(b_{jt}))$. Since c is the same for all bidders, there exists a common thresholds $\underline{b}$ for the bids: While bidders with optimal bids above $\underline{b}$ participate, all the others wait for one of the next auctions where their draw of ϵ is higher.[4] Bidder i can thus compute his winning odds either as the probability of being higher than the highest bid out of the actual competitors, who draw their bids from the parent bid distribution truncated at the participation threshold, or as the probability of being higher than the highest out of all potential competitors, who draw their bids from the full parent bid distribution F_b. The latter viewpoint will be maintained in the following. Let $b_h \equiv b^{n-1:n-1}$ denote the highest bid out of a sample of $n-1$ from the parent F_b. The distribution of this order statistic is then given by $F_b^h \equiv (F_b)^{n-1}$.

Learning. While the model stipulates, the bidder cannot trace a specific competitor or a specific sample of competitors by observing bids and participation decisions, it allows for the possibility that the bidder is uncertain about the underlying primitives, namely the bid

[3]To be fully correct, competitor j's strategy also depends on his information set since he also engages in learning. However, by the same reasoning, it is unrealistic that bidder i can assess how learning on the side of the competitors changes the distribution of bids he faces and he thus believes that it does not affect stationarity.

[4]For a formal proof see Sailer (2005), Lemma **??**.

distribution from which the bids of his competitors are randomly drawn in each auction and that he can learn about it by observing realizations of this distribution. By assumption, only participants in an auction can observe these realizations.[5] Learning is assumed to happen in a Bayesian way, that is, the bidder incorporates new relevant information into his prior beliefs and future decisions rely on the posterior distribution (see e.g. DeGroot (1970)). More specifically, the distribution of the unknown variable b_h is known up to a parameter θ; the bidder does not know the value of this parameter but can specify a prior probability distribution over it. With each observation, that is, after each auction he participated in, he can update this prior with all relevant information he obtained. The result is a posterior distribution of the parameter θ and a new predictive distribution for b_h.

The Bidders' Problem. Taking account of the auction environment and incorporating the learning aspect yields the following expected per period profit function for a participating bidder in period t:

$$\mathbb{E}_{t-1}\left[\mathbf{1}\{b_{it} > b_{ht}\}(v_i + \epsilon_{it} - b_{ht})\right] - c \tag{1}$$

The decision to participate, first of all, involves paying the bidding cost c. If the bidder wins, which is the case when his bid is higher than the highest of his competitors, he gets his valuation and pays the price determined by the highest bid of his competitors.[6] The time subscript on the expectation operator indicates that the bidder can incorporate all information that has realized up to period $t-1$ in his expectation over the unknown variable b_{ht}. A non-participating bidder obtains zero in that period.

Since there is an infinite number of consecutive auctions, the bidder can try again in the next auction whenever he looses. In recursive form, bidder i's intertemporal optimization problem is

[5] Letting also non-participants to an auction learn as soon as they have entered the marketplace would change the problem insofar as the continuation value when the bidder decides not to participate would also be a function of the updated parameters (see equation (2)). See also argumentation in the empirical part.

[6] In the following I will ignore the minimal increment since it is usually very small in comparison with the price to be paid.

thus given by:

$$W_i(\epsilon_i, \theta_i) = \begin{cases} \max\left\{ \max_{b_i \geq 0} \mathbb{E}\left[\mathbf{1}\{b_i > b_h\}(v_i + \epsilon_i - b_h) - c + \mathbf{1}\{b_i \leq b_h\} W_i^e(\theta_i') \,|\theta \right], \, W_i^e(\theta_i) \right\} & \\ & \text{before winning} \\ 0 & \text{after winning} \end{cases} \tag{2}$$

where $W_i^e(\theta_i) \equiv \int W_i(\epsilon_i, \theta_i) f_\epsilon(\epsilon) d\epsilon$ denotes the expected continuation value. Equation (2) states, whenever the bidder participates and does not win, his return is the updated continuation value $W_i^e(\theta_i')$ minus the bidding costs. When the bidder decides not to participate in that auction, he gets the continuation value $W_i^e(\theta_i)$ which does not include any updating of the parameter. In case the bidder wins, his continuation value becomes zero and he leaves the market. I will drop the case $W_i(\cdot) = 0$ as well as the subscript i in the following for notational ease.

Updating in Second Price Auctions. So far nothing has been said about what information the bidder could use for updating. Even if learning is directly over the distribution of the highest bid, it is not straightforward since the information the bidder is primarily interested in, b_{ht}, is not observable in a second price auction. Is there alternative information which the bidder could use instead to improve his prior over the parameter of the distribution of b_h? What the bidder can observe is the full bidding history below the transaction price. He therefore could use some or all of the lower ordered bids. Since these are drawn from the same parent as b_h, there exists a relation between these bids. In the following, I explore whether and how this relation can be exploited for Bayesian updating in the above dynamic optimization problem.

Given the bid information, the bidder has in principle two options. First, he can use all or part of the available lower bids to update the parent distribution. From the parent he then derives the distribution of the order statistic he is interested in. Updating a distribution from the sort of incomplete data encountered here, that is, data that lacks the highest bid and those that fall below the participation thresholds, is though far from easy; incorporating this into a recursive dynamic programming problem seems to be doomed to failure. Directly learning about a parameter of the relevant order statistic distribution appears more straightforward at first glance; given that in general distributions of order statistics are merely functions of the underlying parents, the same parameters would, though, have to be updated unless specific functional forms for the distributions are assumed. The problem of the latter approach is that a convenient form for the distribution of any order statistic, e.g. the normal (see Burdett and Vishwanath (1988)), might lead to an implausible parent. Further, the related distribution of any other order statistic that is used for updating would not necessarily have the same or any other convenient functional form. The computational problems would thus not change. I will show in the following that if one is willing to rely on asymptotic distributions, analytical results can be obtained in a dynamic programming problem without making ad hoc distributional assumptions for the parent. Before continuing, a digression on the asymptotics of order statistics distributions is in order.

Asymptotic Order Statistics Distributions. Since Gumbel (1958), if not before, it is known that if the highest order statistic $b_h = b^{n:n}$ of any parent has, upon suitable standardization, a limiting distribution as $n \to \infty$, this will be one of either Gumbel (double exponential), Weibull, or Frechet type. Gumbel, often also referred to as *the* extreme value distribution, is by far the most common type. Its density is given by:

$$g(z_h) = e^{-e^{-z_h}} e^{-z_h} \text{ with } z_h = (b_h - \theta)\tau. \tag{3}$$

θ and τ denote the location and scale parameters which are used for standardization.[7] While for normal distributions convergence has been found to be excessively slow ($n > 100$), for most other distributions already much smaller sample sizes lead to good approximations.

It has further been shown that not only the highest but also all lower order statistics (k-th extremes, $b_{(k)} \equiv b^{n-k+1:n}$ and $b_{(1)} = b_h$) have asymptotic densities given by:[8]

$$g^{(k)}(z_{(k)}) = k^k([k-1]!)^{-1} e^{-ke^{-z_{(k)}}} e^{-kz_{(k)}} \text{ with } z_{(k)} = (b_{(k)} - \theta)\tau \tag{4}$$

It should be noted that all limiting densities have the same normalizing constants.

Another result, which will prove helpful, is that the joint density of the top k extremes asymptotically converges to:

$$g_{1,\dots,k}(z_{(1)}, \dots, z_{(k)}) = G(z_{(k)}) \Pi_{i=1}^{k} \frac{g(z_{(k)})}{G(z_{(k)})}, \tag{5}$$

[7]The variable $z = (b_h - \theta)\tau$ has mode 0, median $\ln(\ln(2))$, mean γ, and variance $\pi^2/6$. The optimal choice of the standardizing parameters θ and τ depends on the sample size, n.

[8]I only report the Gumbel form. Also here three limiting types are possible where the other two are similar to Weibull and Frechet distributions.

where $g(\cdot)$ is defined as in (3) and $G(\cdot)$ denotes the corresponding cdf. (These and the preceding results as well as further discussion can be found in David and Nagaraja (2003, Ch. 10) and Kotz and Nadarajah (2002).) In the following, a tilde depicts the density of the bid corresponding to g: $\widetilde{g}(b;\theta,\tau) = \tau g((b-\theta)\tau)$. Combining these results the joint density of b_h and $b_{(2)}$ is given by:

$$\widetilde{g}_{1,2}(b_h, b_{(2)};\theta,\tau) = \tau^2 \frac{g((b_h-\theta)\tau)g((b_{(2)}-\theta)\tau)}{G((b_h-\theta)\tau)} = \tau^2 e^{-(b_h-\theta)\tau} g((b_{(2)}-\theta)\tau)). \tag{6}$$

As mentioned above, uncertainty will be defined as uncertainty about a parameter of the distribution of b_h. Let this be the location parameter θ. The bidder has a subjective prior over this parameter which is distributed according to $\xi_{\text{prior}}(\theta)$. With each participation the bidder incorporates new information in a Bayesian way. If the information he uses is the k-th order statistic, the posterior distribution, that is, the distribution of the parameter conditional on the observation $b_{(k)}$, is computed from:[9]

$$\xi_{\text{post}}\ (\theta|b_{(k)}) = \frac{\tilde{g}^{(k)}(b_{(k)}|\theta)\xi_{\text{prior}}(\theta)}{\int \tilde{g}^{(k)}(b_{(k)}|\theta)\xi_{\text{prior}}(\theta)d\theta}$$

If the prior and the posterior distribution belong to the same family, that is, they differ only by the value of a finite parameter vector, we say that ξ and $g^{(k)}$ constitute a conjugate family. Using conjugate priors facilitates the updating procedure since at each step only the parameters of ξ change as a function of the observed information but not the distribution itself.

When the scale parameter τ is known, all of the asymptotic extreme value distributions belong to the exponential family. For members of the exponential family there always exists a sufficient statistic of fixed dimension (see e.g. DeGroot (1970, Ch.9)). This statistic is given by $t(\mathbf{b}_{(k)}, N) = \sum_{i=1}^{N} e^{-\tau b_{(k),i}}$,[10] where N denotes the number of independent observations of the same k-th order statistic which are used for updating, and the vector $\mathbf{b}_{(k)}$ collects all the observations. On the other hand, if such a statistic exists then there also exists a conjugate family for this distribution:

Lemma 1. *A conjugate family for the Gumbel distribution is given by:*

$$\xi^{(k)}(\theta; r, C) = \frac{\tau\ \left(kre^{\theta\tau}\right)^{kC} e^{-kre^{\theta\tau}}}{\Gamma\left(kC\right)}, \tag{7}$$

[9] So far I used the notation $f(x;\theta)$ to denote the density of x for each $\theta \in \Theta$. To highlight that θ now is the value of a random variable I will in the following denote $f(x|\theta)$ as the density conditional on a specific realization of θ.

[10] A consistent estimator for θ given τ is given by $\widehat{\theta} = -\tau^{-1} ln\left(\frac{1}{N}\sum_{i=1}^{N} e^{-b_h\tau}\right)$ (see Kotz and Nadarajah (2002)).

where $x = e^{\theta\tau}$ is distributed according to a gamma distribution with shape parameter kC ($C \geq 1$)[11] *and scale parameter $1/(kr)$.*[12] *C and r are the parameters of this distribution which are updated with each new observation as follows:*

$$r_k' = r_k + t(b_{n-k+1}, 1) \text{ and } C' = C + 1. \tag{8}$$

τ, the scale parameter of the Gumbel distribution, is known and thus not part of the updating process. k is the order of the statistic that is used for updating.

Proof. See Appendix. □

To build the expectation in (2) over b_h we first have to integrate out the unknown parameter θ. The resulting density of b_h is parameterized by the parameters of the posterior distribution ξ and is called the predictive distribution. The following lemma shows that under the Gumbel assumption this density has a well known form as well:

Lemma 2. *The predictive distribution of b_h when b_h is known to have a Gumbel distribution with scale parameter τ is given by:*

$$h(b_h; \bar{r}, C) = \frac{1}{\bar{r}} \left(\frac{\bar{r}}{\bar{r} + (kC)^{-1} e^{-\tau b_h}} \right)^{kC+1} \tau e^{-\tau b_h} \tag{9}$$

where $x = e^{-\tau b_h}$ is distributed according to a Generalized Pareto distribution (GPD) with scaling parameter $\bar{r} = r/C$ and shape parameter $1/kC$.[13]

Proof. See appendix. □

The Bidder's Problem Continued. The observation most commonly available in auction data sets is the transaction price which corresponds to the second highest bid. The following analysis thus concentrates on the case when this statistic is used for updating. The preceding results can then be combined to restate the bidder's problem as follows:

[11] This assumption guarantees that the predictive distribution has finite mean and variance for whatever statistic is used for updating, including $k = 1$.

[12] The expected value of the gamma distribution is $\mathbb{E}(x) = \frac{C}{r}$, the expected value of θ is given by $\mathbb{E}(\theta) = \frac{\Psi(2C) - ln(2r)}{\tau}$, where Ψ denotes the digamma function.

[13] The GPD is the limiting distribution for extreme exceedances. The threshold for x here is 0, that is this GPD gives the exceedances of x above 0. For $1/2C = 0$ the distribution is equivalent to the Gumbel distribution, for $1/2C > 0$ the tails are heavier. For $\bar{r} = 1$ the distribution is also called *Generalized Standard Pareto Distribution.*

Proposition 1. *Assuming the number of potential participants in an auction is large enough to apply the asymptotic results discussed in the previous paragraphs, the bidder's problem when the transaction price is used for updating is given by:*

$$W(\epsilon, \bar{r}, C) = \max\left\{\max_{b\geq 0} \int_{-\infty}^{b} (v + \epsilon - b_h) h(b_h; \bar{r}, C) db_h - c\right) + \int_{b}^{\infty} \int_{-\infty}^{b_h} W^e(\bar{r}', C') \cdot$$
$$\cdot\, h(b_{(2)}; \bar{r}, C) i(b_h, b_{(2)}; \bar{r}, C) db_{(2)} h(b_h; \bar{r}, C) db_h, W_i^e(\bar{r}, C)\right\}$$
$$s.t.\ C' = C + 1, \bar{r}' = \bar{r}\frac{2C}{2C+1} + \frac{1}{2C+1} e^{-\tau b_{(2)}} \text{ and } r_0, C_0 \text{ given,}$$

where

$$i(b_h, b_{(2)}; \bar{r}, C) = \frac{2C+1}{2C} \frac{\bar{r}}{\bar{r} + (2C)^{-1} e^{-\tau b_{(2)}}} \left(\frac{\bar{r} + (2C)^{-1} e^{-\tau b_h}}{\bar{r}}\right)^{2C+1}$$

and $h(\cdot)$ as defined in Lemma 2.

Proof. See Appendix. □

The advantage of this formulation as opposed to the problem stated in (2) is that it uses analytic solutions for the distributions which are used for expectation building without significantly restricting the parent distributions from which the bids are drawn. Further, it gives the bidder clear advice how to incorporate new information in form of transaction prices into his beliefs.

To be able to provide analytical results for the optimal strategies, the impact of the unknown variable $b_{(2)}$ on the value function has to be disentangled from the impact of the other variables. A first step forward is to separate the unknown b_h into a function of the stochastically evolving scaling parameter $\bar{r}$ and a random variable whose realization is independent of $\bar{r}$. From this "standardized" version of the problem (see Appendix) the following guess for the value function emanates which isolates the influence of the stochastic variable $b_{(2)}$:

$$W(\epsilon, \bar{r}, C) = \frac{ln(\bar{r})}{\tau} + \alpha(\epsilon, C). \tag{10}$$

$\alpha(\epsilon, C) \equiv W(\epsilon, 1, C)$ denotes the part of the value function which evolves deterministically with C.

Under this guess, it is possible to derive analytic forms for the optimal strategies. These are given in Proposition 2. The proposition also states that the guess for the value function is correct. All details of the computation are provided in the Appendix.

Proposition 2.

(a) The value function can be written in the form:

$$W(\epsilon, \bar{r}, C) = \frac{ln(\bar{r})}{\tau} + \alpha(\epsilon, C) \tag{11}$$

with $\alpha(\epsilon, C) \equiv W(\epsilon, 1, C)$, $W^e(\bar{r}, C) = \frac{ln(\bar{r})}{\tau} + \alpha^e(C)$, *and* $\alpha^e(C) \equiv \int \alpha(\epsilon, C) dF_\epsilon(\epsilon)$.

(b) The bidder's optimal bid is given by:

$$b^* = b(\epsilon, \bar{r}, C) = v + \epsilon - \frac{\ln(\bar{r})}{\tau} - \alpha^e(C+1) - l(\epsilon, C). \tag{12}$$

with

$$l(\epsilon, C) \equiv \frac{1}{\tau}\left(\frac{1}{2C+1} + \ln\left[\frac{2C}{2C+1-e^{\frac{1}{2C+1}+\tau(\alpha^e(C+1)-v-\epsilon)}}\right]\right).$$

Proof. See Appendix. □

The optimal bid, thus, also separates into a part which changes with the new prior for $\bar{r}$ and another part which deterministically evolves with C. This part consists of $\alpha^e(C+1)$ and an additional component $l(\epsilon, C)$ which vanishes with $C \to \infty$.[14]

The first term of $l(\epsilon, C)$ clearly is positive. The sign of the second term is less obvious. If $\alpha^e(C+1) > v + \epsilon$[15] then the numerator is bigger than the denominator and the logarithmic term in $l(\cdot)$ is positive. Due to the learning possibilities, the bidder would hence shade his bid.

When no learning happens anymore ($C = \infty$), the optimal bidding strategy collapses into one where the bidder shades his valuation by his continuation value:

$$b^*_{C\to\infty} = v + \epsilon - \frac{\ln(\bar{r})}{\tau} - \alpha^e(\infty) = v + \epsilon - W_i^e(\bar{r}, \infty). \tag{13}$$

This is exactly the same solution as in the case where the bidder knows the competitors' bid distribution from the beginning and learning is thus not necessary (see Sailer (2005)).

[14]This is true as long as $\alpha^e(C+1)$ does not increase "too quickly" with C. In the next section I show by numerical simulation that $\alpha^e(C+1)$ actually decreases in C for reasonable values of the parameters.

[15]The simulation exercise shows that this is true for reasonable values of the parameters.

3 Impact of Learning on Optimal Bidding Strategies

Learning influences the bidder's optimal bidding behavior in two ways: First, the prior over the distribution of the location parameter changes with each new observation.

Proposition 3. *The optimal bid increases with the observed transaction price; the bid increases less the higher C.*

Proof. $\frac{\partial b^*}{\partial b_{(2)}} = \frac{\partial b^*}{\partial \bar{r}} \frac{\partial \bar{r}}{\partial b_{(2)}} = \frac{1}{\bar{r}C} e^{-\tau b_{(2)}}$ and $\frac{1}{\bar{r}(C+1)} e^{-\tau b_{(2)}} - \frac{1}{\bar{r}C} e^{-\tau b_{(2)}} < 0.$ $\square$

In a second price auction any influence of the competitive environment can work only via the continuation value. The higher the last observed transaction price was, the lower the prior of $\bar{r}$ and, thus, the lower the winning chances in the future. Lower future expected returns decrease the continuation value. Since the bidding strategy involves shading the valuation by the continuation value, the bidder will now shade less and the bid is thus higher. The effect becomes less pronounced with each participation since the more observations the bidder has already sampled the less influential is any new information.

The influence of the number of observations works via the same channel. The results are, however, less clear cut. Differencing the optimal bid with respect to C gives:

$$\begin{aligned} b(\epsilon, \bar{r}, C+1) - b(\epsilon, \bar{r}, C) = \alpha^e(C+2) - \alpha^e(C+1) + \frac{1}{\tau} \frac{2}{(2C+1)(2C+3)} + \\ + \frac{1}{\tau} \ln \left[\frac{2C^2 + 3C - Ce^{\frac{1}{2C+3} + \tau(\alpha^e(C+2) - v - \epsilon)}}{2C^2 + 3C - (C+1)e^{\frac{1}{2C+1} + \tau(\alpha^e(C+1) - v - \epsilon)} + 1} \right] \end{aligned} \tag{14}$$

A sufficient condition for this difference to be positive is $\alpha(C+2) - \alpha(C+1) < \frac{1}{\tau} \frac{2}{(2C+1)(2C+3)}$.

To be able to say more about the evolution of bids during the course of learning α thus has to be characterized more closely. Combining the standardized form of the value function with the value function guess provides a form for α (see proof of Prop. 2). Solving the integrals as far as possible and rearranging gives a functional form which separates the total value into a value from winning the object, which is similar to the one in the problem without learning, and a value from learning:

$$\alpha(\epsilon, C) = \max\Bigg\{ \underbrace{\int_{b^{0*}}^{\infty} \left(v + \epsilon + \frac{\ln(x)}{\tau}\right) h_0(x;C)dx}_{\text{Ex. value of winning } (ER_w(\epsilon,C))} - c + \int_{0}^{b^{0*}} \alpha^e(C+1) h_0(x;,C)dx$$

$$\underbrace{+\frac{\frac{4C+1}{(2C+1)2C} - \ln\left(\frac{2C+1}{2C}\right)}{(2C)^{2C}} - \frac{\ln\left(\frac{2C+b^{0*}}{2C+1}\right) + \frac{4C+1}{(2C+1)2C}}{(2C+b^{0*})^{2C}}}_{\text{Ex. value from learning } (ER_l(\epsilon,C))}, \alpha^e(C) \Bigg\} \tag{15}$$

with $h_0(x;C) = \left(\frac{2C}{2C+x}\right)^{2C+1}$

and $b^{0*} = b^0(\epsilon, C) = \frac{1}{\bar{r}} e^{-\tau b^*} = 2C \left(e^{-\tau(\alpha^e(C+1)-v-\epsilon) - \frac{1}{(2C+1)}} (2C+1) - 1 \right)^{-1}$.

There are thus two influences of an increase in C on the bidder's value: Clearly, $ER_l(\epsilon, C)$ decreases in C and eventually vanishes. Diminishing returns from learning are also consistent with intuition. The second influence comes via the predictive distribution $h_0(x;C)$ which the bidder uses for determining his expected value of winning. With each new observation this distribution becomes less dispersed. Differencing the expected value of winning with respect to C and using integration by parts gives:

$$\begin{aligned} ER_w(\epsilon, C)) - ER_w(\epsilon, C+1)) = &- \int_{b^0(\epsilon,C+1)}^{b^0(\epsilon,C)} \left(v + \epsilon + \frac{\ln(x)}{\tau}\right) h(x; C+1)dx + \\ &+ \left(v + \epsilon + \frac{\ln(b^0(\epsilon, C))}{\tau}\right) \left(H_0(b^0(\epsilon, C); C+1) - H_0(b^0(\epsilon, C); C)\right) \\ &+ \int_{b^0(\epsilon,C)}^{\infty} \left(v + \epsilon + \frac{\ln(x)}{\tau}\right) \left(H_0(x; C+1) - H_0(x; C)\right) dx \end{aligned} \tag{16}$$

The last two lines are positive since $H_0(x;C) < H_0(x;C+1)$ for all x. This result shows that the bidder prefers to play against a more dispersed distribution of competitors' bids. On the other hand, the bidder is free to choose the cutoff values b^{0*} in each period (first line). By choosing a lower bid in $C+1$ than in C the bidder can counteract the negative effect. The overall effect is ambiguous.

Given this uncertainty over how $ER_w(C)$ changes with C, it is not possible to determine analytically how α^e and thus the bids change with each new trial. I therefore compute the value function numerically for values of the parameters which match the latter empirical analysis. The following figure shows the results for α^e and $\mathbb{E}[b^* + ln(\bar{r})/\tau]$. [16]

[16]Since I do not want to pick out an arbitrary realization of ϵ, I here report the mean. Bids also increase for all possible realization of ϵ.

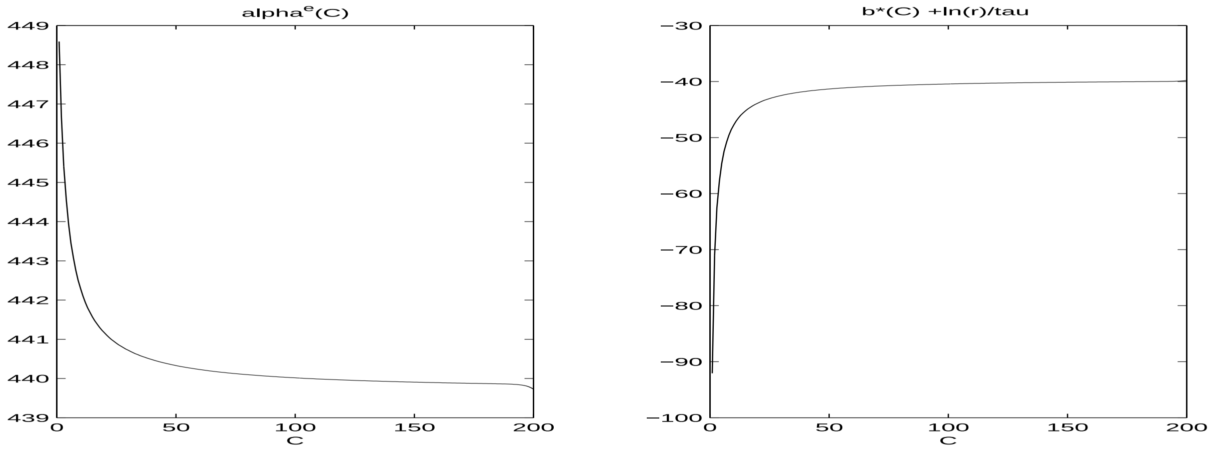

Figure 1: Simulation results for α^e and $\mathbb{E}[b^* + ln(\bar{r})/\tau]$

($v = 400$, $c = 8$, $f_\epsilon(\epsilon) = \mathcal{N}(0, 10)$, $\tau = 0.016$, and $C_0 = 2$. C is assumed to be close to infinity at $C = 200$)

For the given parameter values α^e thus decreases and the optimal bid given any value of $\bar{r}$ increases in C. The increases however quickly become smaller and eventually vanish.

Proposition 4. *If $\alpha^e(C)$ decreases in C the optimal bid increases in C. This is for example true for $v = 400$, $c = 8$, $f_\epsilon(\epsilon) = \mathcal{N}(0, 10)$, $\tau = 0.016$, and $C_0 = 2$.*

Proof. Follows from the preceding discussion and the simulation exercise. □

4 Detecting Learning in eBay Data

A characterizing feature or Internet auctions is that in many product categories there are a large number of buyers and sellers present, much like in a marketplace, which makes the search framework of the benchmark model applicable.

eBay is in most countries the major auction platform. I collected data on a Personal Digital Assistant (PDA or palm pilot), the Compaq Ipaq H3850, from eBay Germany over a 7 month period in 2002. This data is in the following used to test the implications of the learning model. Every 4 1/2 hours an auction for this product closed. The average price bidders paid for it was 477€. 3829 distinct bidders are observed in the sample who on average placed bids in 2 different auctions. Many, however, bid many more times. (For a more detailed description of the data set and the variables used in the latter estimation as well as a frequency distribution for the number of bids per bidder see Sailer (2005).).[17]

Does the eBay setting conform with the main assumptions of the benchmark model? First of all, the auction rules at eBay are a mixture of an open ascending and a sealed bid second price auction. The literature on eBay, though, argues that the Vickery assumption presents a good approximation to the true rules since it is not optimal for a bidder to reveal private information early on. The data confirms this; bidding activity is concentrated at the end of an auction.

An estimate of the potential number of bidders is obtained when adding to the actual bidders in an auction those that bid in any previous auction without winning and show up later again in another auction.[18] The data reveals, on average around 140 bidders are looking for the Compaq 3850 in any given auction. Further, there are probably bidders who never show up in my sample at all but are still considering to bid for the product. Using asymptotic results, thus, is justified for this product category.

As opposed to eBay.com, at eBay.de it was not possible to search for past, closed auctions at the time the data was collected. While the potential bidder can also use eBay's function "observe auctions" to learn about his competitors' bids before starting to bid himself, he would have to spend considerable time to figure out what the average transaction price is, given that an average eBay auction lasts 5 days. This is costly. An alternative is to participate right away with a low bid using an initial guess for the competitors' bid distribution. As opposed to merely observing, this strategy involves not only a cost but also a winning chance. If the bidder actually learns from past observations in the way described by the model, there is thus a good chance that this shows up in the data.

Bidding in the model is assumed to be costly. eBay does not charge the bidder any fee; however, there is a cost in terms of the time spent in front of the computer and the connection charges. Given this cost, the bidder has to take a participation decision in each auction. The bidder participates if his draw of ϵ is so high that it leads to an optimal bid above $\underline{b}$. The question is whether, once the bidder has started bidding for a product, he would also observe

[17] Here, I use the full bidding data, if not otherwise mentioned, as opposed to Sailer (2005) where only those bids submitted in the last 10% of the auction time entered the estimation. The reason is that low bids now might actually represent optimal strategies. In the latter estimation I will present results for both, the full data set and the data which only includes bids submitted in the last 10% of the time.

[18] See also the description of the participation vector I_{P_S} in the Appendix to Sailer (2005).

the transaction prices in those auctions that he decided not to participate in? While in principal he could, it is hard to judge which of the auctions he actually observed since there is no click data available. In the following I will therefore restrict attention to those auctions in which the bidder is observed with a bid.

Finally, the model assumed an infinite sequence of identical products. At eBay there are hardly any two products that are exactly the same. This is also true for the Compaq 3850. Many of them come with extras or are used, some have defects or a foreign operating system. To account for the heterogeneity in the estimation, bidders' valuations for products have to be adjusted. I assume that bidder i's valuation for a product is made up out of an individual part v_i, the weighted sum of its k=1,...,K product characteristics, $\mathbf{x} = (x_1, \ldots, x_K)$, with weights common to all bidders, and an individual and auction specific error term ϵ_{it}:

$$v_{it} = v_i + \mathbf{x}_t\beta + \epsilon_{it}.$$

From the model and the simulation exercise we know how bidders should react to new information in form of past observed transaction prices and with each new trial. This behavior will be captured by the following reduced form for the bid:

$$b_{it} \cong v_i + x_t\beta + \epsilon_{it} - W_i^{\text{red}}(\mathbf{b}_{(2)}^{o,it}, C_{it}) \tag{17}$$

where the vector $\mathbf{b}_{(2)}^{o,it}$ collects all observations of transaction prices which bidder i has made up to auction t, starting with the first one, and C_{it} is equivalent to the number of previous trials. Since the influence of the number of trials and the past observations of transaction prices is via the continuation value, which is used to shade the valuation, the learning component W_i^{red} enters the bid equation negatively. It is specified as follows:

$$W_i^{\text{red}} = a_{i0} + a_1 C_{it} + a_2 C_{it}^2 + \mathbf{b}_{(2)}^{o,it}\gamma_{it}, \tag{18}$$

with $\gamma_{it} = (\gamma_1, \ldots \gamma_{t_i}, \ldots \gamma_{C_{it}})$. From the model we know that a_1 and a_2 should be negative and positive, respectively, that is, the continuation value decreases with each new trial but in a decreasing way over time. All the γ-coefficients should be negative since the future winning chances decrease the higher the transaction price observations. Since the influence vanishes over time, the coefficients should be smaller the higher the index.

Data from eBay auctions does not represent a random sample from the parent bid distribution due to missing winning bids and the participation decision. (See Sailer (2005) for a detailed discussion.) The estimation results in Sailer (2005) show that estimates obtained by simple first differences, however, provide good approximations to the true coefficients. I therefore estimate the coefficients by OLS from:

$$b_{it} - b_{i,t-1} = (\mathbf{x}_t - \mathbf{x}_{t-1})\beta + a_1 + a_2\left((C_{it})^2 - (C_{i,t-1})^2\right) + \left(\mathbf{b}_{(2)}^{o,it} - \mathbf{b}_{(2)}^{o,it-1}\right)\gamma_{it} + \epsilon_{it} - \epsilon_{i,t-1}. \tag{19}$$

Note that due to the differencing only the last observed transaction price actually enters this equation.

The transaction prices also reflect the product features. The bidder has to account for this heterogeneity. Given the bidder knows $\mathbf{x}_t\beta$, the news in each observation of a transaction price is the amount the second highest bid is above this common component. In Specification (1) I ignore this fact and use the bids as observed. In Specification (2) the observed transaction prices are first homogenized using coefficients from a simple OLS regression of product characteristics onto bids: Each element in $\mathbf{b}_{(2)}^{o,it}$ is replaced by $b_{(2),t_i}^{o,it} + (\mathbf{x}_t - \mathbf{x}_{t_i})\beta_{OLS}$.

Since it is possible to argue that only those bids submitted towards the end represent optimal strategies, I repeat the estimation using only bids submitted in the last 10% of an auction (Specification (3)). Past observations of transaction prices are still those that were observed in all auctions a bidder participated in - also when his bid in that auction was submitted early on - and are used as observed, that is, without prior homogenization.

Table 1 reports the results for the different specifications. First of all, there is a pronounced negative time trend in the data. This is attributed to the high tech characteristic of the product and also observed in shops outside eBay. The signs of the coefficients for product characteristics are, besides the variable CAREPAQ, all significant and have the expected signs: The age, a foreign operating system, and a defect have a negative impact on the bids while additional extras lead to higher bids. The relative importance of the different extras matches their relative prices outside eBay.

Table 1: Estimates of the Learning Model Using a Linear Specification for the Bids.

	(1)		**(2)**		**(3)**	
TREND	- .984	(.218)	- .970	(.218)	- .460	(.119)
AGE	- .031	(.022)	- .031	(.022)	- .065	(.017)
AGE_NS	-9.224	(4.906)	-9.289	(4.906)	-11.377	(3.525)
DEFECT2	-39.886	(12.864)	-39.711	(12.864)	-31.869	(11.270)
OS_ENG	-10.515	(8.858)	-10.492	(8.858)	-18.469	(6.942)
OS_FRENCH	-202.630	(37.047)	-202.510	(36.047)	-77.345	(16.738)
EXTRAS	5.469	(3.270)	5.448	(3.270)	6.338	(2.663)
JACKET1	26.364	(16.786)	26.560	(16.786)	18.367	(8.655)
JACKET5	102.570	(22.189)	102.130	(22.190)	113.380	(25.612)
MEMORY	.157	(.051)	0.156	(.051)	.227	(.043)
HARDDISK	69.800	(16.674)	69.541	(16.674)	80.582	(22.346)
NAVIGATION	73.952	(21.968)	74.068	(22.968)	184.960	(29.883)
CAREPAQ	- .699	(5.618)	- .704	(5.618)	10.542	(4.946)
a_1	-9.114	(5.315)	-9.213	(5.315)	-11.009	(4.189)
a_2	0.614	(.493)	.618	(.493)	.966	(.415)
γ_1	- .039	(.017)	- .032	(.017)	- .015	(.010)
γ_2	- .010	(.012)	- .007	(.012)	.001	(.007)
γ_3	- .001	(.013)	.000	(.013)	- .002	(.008)
OBS	2479		2479		858	
R^2	.129		.129		.407	
adj R^2	.123		.123		.395	

White heteroscedasticity robust estimation. Standard errors in parenthesis.

The coefficients which measure the influence of the past number of trials, a_1 and a_2, are negative and positive, respectively, that is, they conform to the predictions of the model. Both coefficients are significant at the 10% level in all specifications. As for the influence of the past observations, reflected in the γ coefficients, the results are less pronounced. The maximum number of observations of transaction prices taken by an individual, that is, the maximum number of auctions he already participated in, is larger than 16. Since I do not have enough observations, I cannot estimate all of these coefficients. Instead, I restrict attention to the first 3 observations a bidder took and see whether they fulfil the predictions of the theoretical model. The coefficient for the first observation, γ_1, is significantly negative in all specifications and, thus, in line with the model. While mostly the influence of the second and third observation is negative as well and smaller, the effects are insignificant. A reason for this insignificance might be that part of the learning is not reflected in the data, that is, the bidder also learns when he does not participate.

There is nearly no difference in the results between Specification (1) and (2). This is not too astonishing since on average the influence of high and low value products cancels. There are some small differences when using Specification (3). These could, however, also be attributed to the considerably smaller sample size. The problem with first differences in the eBay setting is that half of the bidders are lost since they only participate once. This is reflected in the relatively low R^2. This effect is even more pronounced when only using bids submitted in the last 10% of the time. If the one time bidders are different from the other bidders (see Sailer (2005)) this would bias the results.

5 Conclusion

Standard sequential auction models cannot rationalize increasing individual bids which are encountered at eBay. The paper showed that uncertainty over the competitors' primitives can provoke such bidding patterns. By augmenting the search model proposed in Sailer (2005) by a Bayesian learning procedure it was not only shown that the bidder optimally increases his bids over time but also how he will incorporate new information in form of past observations of transaction prices into his bidding strategy. Results from reduced form estimations gave evidence that learning is a possible explanation for the observed bidding patterns.

Besides showing that there is evidence of Bayesian updating in the eBay data, a main contribution of this work is to show how to make use of asymptotic order statistic distributions in an auction environment. Learning in second price auctions is complicated by the fact that the bidder cannot directly observe the statistic of interest, namely the highest bid of his competitors, but only the transaction price, which is the second highest bid. Using results from the asymptotic distribution of extreme order statistics, it was shown that the transaction price can also be used to update the beliefs over a parameter of the distribution of the highest bid.

While the empirical part allowed for bidder heterogeneity in form of different valuations, all bidders started with the same priors. It would be interesting to see whether the results are robust to heterogenous priors. Since such a specification would require a more structural approach, this has to be deferred to future research. A structural approach would also be desirable to see whether the results are robust to possible non-linearities.

Finally, the results hinge on the premise that the proposed model is the true model. While increasing bids over time and with new observations of transaction prices intuitively make sense, derivation of the optimal bidding strategy is far from easy. It would be interesting to see whether other bidding strategies, including e.g. reasonable heuristics or behavioral strategies, would explain the data better.

Appendix

6 Proofs

Proof of Lemma 1. The posterior distribution of the parameter, given that only one observation is included at a time which is distributed Gumbel, is given by:

$$\begin{aligned}\xi_{\text{post}}(\theta|b_{n-k+1}) &= \frac{\tilde{g}^{(k)}(b_{n-k+1}|\theta)\xi^{(k)}(\theta;r,C)}{\int_{-\infty}^{\infty}\tilde{g}^{(k)}(b_{n-k+1}|\theta)\xi^{(k)}(\theta;r,C)d\theta} \\ &= \frac{e^{-ke^{-(b_{(k)}-\theta)\tau}}e^{-k(b_{(k)}-\theta)\tau}e^{\theta\tau kC}e^{-kre^{\theta k}}}{\int_{-\infty}^{\infty}e^{-ke^{-(b_{(k)}-\theta)\tau}}e^{-k(b_{(k)}-\theta)\tau}e^{\theta\tau kC}e^{-kre^{\theta k}}d\theta} \\ &= \frac{e^{\theta\tau k(C+1)}e^{-e^{\theta\tau}k\left(r+e^{-b_{(k)}\tau}\right)}}{\int_{-\infty}^{\infty}e^{\theta\tau k(C+1)}e^{-e^{\theta\tau}k\left(r+e^{-b_{(k)}\tau}\right)}d\theta}.\end{aligned}$$

It is easily shown that the denominator integrates to $\Gamma\left(k(C+1)\right)/\left(\tau\left(k\left(r+e^{-b_{(k)}\tau}\right)\right)^{k(C+1)}\right)$ (Use equation 7 and the fact that distributions integrate to 1) and thus:

$$\xi_{\text{post}}(\theta|b_{n-k+1}) = \frac{\tau\left(k(r_k+t(b_{n-k+1},1))e^{\theta\tau}\right)^{k(C+1)}e^{-k(r_k+t(b_{n-k+1},1))e^{\theta\tau}}}{\Gamma\left(k(C+1)\right)}. \tag{20}$$

Letting $r'_k = r_k + t(b_{n-k+1},1)$ and $C' = C+1$ we have $\xi_{\text{post}}(\theta|b_{n-k+1}) = \xi^{(k)}(\theta;r',C')$. □

Proof of Lemma 2. The predictive distribution is given by $\int_\Theta \tilde{g}(b_h|\theta)\xi^{(2)}(\theta;r,C)d\theta$. Using the Gumbel density of the bid and (7) gives:

$$h(b_h;\bar{r},C) = \int_{-\infty}^{\infty}\frac{\tau^2 e^{-e^{-(b_h-\theta)\tau}-(b_h-\theta)\tau}\left(rk\,e^{\theta\,\tau}\right)^{k\,C}e^{-rk\,e^{\theta\,\tau}}}{\Gamma\left(k\,C\right)}d\theta.$$

Now take the constants as far as possible out of the integral, change the variable of integration from θ to $x = e^{\theta\tau}$, and use the fact that $\int_0^\infty e^{-xa1}x^{a2}dx = \frac{\Gamma(a2)a2}{a1^{a2}a1}$ for any positive constants $a1$ and $a2$. It then follows that:

$$h(b_h;\bar{r},C) = \tau\frac{(rk)^{kC}}{\Gamma(Ck)}e^{-\tau b_h}\frac{\Gamma\left(kC\right)kC}{(e^{-\tau b_h}+rk)^{kC+1}}$$

Rearranging finally gives:

$$h(b_h;\bar{r},C) = \tau\left(rk\right)^{kC}kCe^{-\tau b_h}(e^{-\tau b_h}+rk)^{-(kC+1)},$$

which is equivalent to a Generalized Pareto Distribution (GPD). □

Proof of Prop. 1. Using the transaction price for updating in (2) we obtain:

$$
\begin{aligned}
W(\epsilon,\theta) = \max\Bigg\{ & \max_{b\geq 0} \int_{B_h} \mathbf{1}\{b > b_h\}(v+\epsilon-b_h) \int_{\Theta} l(b_h,\theta) d\theta db_h - c + \\
& + \int_{B_h} \mathbf{1}\{b > b_h\} \int_{B_{(2)}} \mathbf{1}\{b_{(2)} < b_h\} W^e(\theta(b_{(2)})) \cdot \\
& \cdot \int_{\Theta} l(b_h, b_{(2)}, \theta) d\theta db_h db_{(2)}, W^e(\theta) \Bigg\}
\end{aligned}
$$

where B_h and $B_{(2)}$ are the set of all possible values of the highest respectively second highest bid and $l(\cdot,\ldots,\cdot)$ denotes the relevant joint distribution of its arguments. Combining this with Lemma 2 for the predictive distribution and using (6) for the joint distribution gives:

$$
\begin{aligned}
W(\epsilon,r,C) = \max\Bigg\{ & \max_{b\geq 0} \int_{-\infty}^{b} (v+\epsilon-b_h) h(b_h;r,C) db_h - c + \\
& + \int_{b}^{\infty} \int_{-\infty}^{b_h} W^e(r',C') \int_{\Theta} \tau^2 e^{-(b_h-\theta)\tau} g((b_{(2)}-\theta)\tau)) \cdot \\
& \cdot \xi(\theta;r,C) d\theta db_h db_{(2)}, W^e(r,C) \Bigg\}.
\end{aligned}
$$

Integrating $\tau^2 e^{-(b_h-\theta)\tau} g((b_{(2)}-\theta)\tau))\xi(\theta;r,C)$ over θ finally gives
$h(b_{(2)};r,C)\frac{2C+1}{2C}\frac{\bar{r}}{\bar{r}+(2C)^{-1}e^{-\tau b_{(2)}}}\left(\frac{\bar{r}+(2C)^{-1}e^{-\tau b_h}}{\bar{r}}\right)^{2C+1} h(b_h;r,C)$. □

Proof of Prop. 2.

I. Standardization.

Upon changing the variables of integration from b_h to $b_h^0 = e^{-\tau b_h}\bar{r}^{-1}$ and from $b_{(2)}$ to $b_{(2)}^0 = e^{-\tau b_{(2)}}\bar{r}^{-1}$ and rearranging we obtain:

$$
\begin{aligned}
W(\epsilon,\bar{r},C) = max\Bigg\{ & \underset{b\geq 0}{max}\left(v+\epsilon+\frac{\ln(\bar{r})}{\tau}\right)(1-H_0(b^0;C)) + \frac{1}{\tau}\int_{b^0}^{\infty} \ln(x)\, h_0(x;C)dx - c + \\
& + \int_0^{b^0}\int_x^{\infty} W^e\left(\frac{2C+y}{2C+1}\bar{r}, C+1\right) h_0(y;C) i_0(x,y;C) dy h_0(x;C) dx, W^e(r,C)\Bigg\}. \quad (21)
\end{aligned}
$$

with $b^0 = e^{-\tau b}\bar{r}^{-1}$, $h_0(x;C) = \left(\frac{2C}{2C+x}\right)^{2C+1}$ with $H_0(x;C) = 1 - \left(\frac{2C}{2C+x}\right)^{2C}$, and $i_0(x,y;C) = \frac{2C+1}{2C+y}\left(\frac{2C+x}{2C}\right)^{2C+1}$. In the following the expression $h_0(y;C)i_0(x,y;C)h_0(x;C)$ will be used in its more concise form $\frac{2C+1}{2C}\left(\frac{2C}{2C+y}\right)^{2C+2}$.

The functions which are used for expectation building are now independent of the stochastic updating parameter $\bar{r}$. Analytical results for the optimal strategies can still not be obtained unless we know the form of the value function as separate functions of the stochastically changing $\bar{r}$ and the deterministically evolving C.

II. Value Function Guess.

Start with the following guess:

$$W(\epsilon, \bar{r}, C) = \frac{ln(\bar{r})}{\tau} + \alpha(\epsilon, C)$$

where $\alpha(\epsilon, C) \equiv W_i(\epsilon, 1, C)$ denotes the mean corrected form of the value function which only depends on C.

Replacing $W^e(\bar{r}', C')$ in (21) by this guess and solving the integrals as far as possible gives:

$$\begin{aligned}
W(\epsilon, \bar{r}, C) = \max \Bigg\{ &\max_{b \geq 0} \frac{\ln(\bar{r})}{\tau} + (v + \epsilon)\left(1 - H_0\left(b^0; C\right)\right) + \frac{1}{\tau} \int_{b^0}^{\infty} \ln(x)\, h_0(x; C) dx - c + \\
&+ \left(\alpha^e(C+1) + \frac{4C+1}{2C(2C+1)\tau} - \frac{1}{\tau} \ln\left(\frac{2C+1}{2C}\right)\right) H_0\left(b^0; C\right) - \\
&- \frac{1}{\tau} \ln\left(\frac{2C + b^0}{2C}\right) \left(1 - H_0\left(b^0; C\right)\right), W^e(r, C) \Bigg\}.
\end{aligned}$$

Under the guess the value function thus separates additively into $\frac{\ln(\bar{r})}{\tau}$ and some other part which only depends on C and b^0 and which has the form:

$$\begin{aligned}
\alpha(\epsilon, C) &= \max \Bigg\{ \max_{b \geq 0} \int_{b^0}^{\infty} \left(v + \epsilon + \frac{\ln(x)}{\tau}\right) h_0(x; C) dx - c + \int_0^{b^0} \alpha^e(C+1) h_0(x; C) dx \\
&\quad + \int_0^{b^0} \int_x^{\infty} \frac{1}{\tau} \ln\left(\frac{2C+y}{2C+1}\right) \frac{2C+1}{2C} \left(\frac{2C}{2C+y}\right)^{2C+2} dy dx, \alpha^e(C) \Bigg\} \\
&= \max \Bigg\{ \max_{b \geq 0} \int_{b^0}^{\infty} \left(v + \epsilon + \frac{\ln(x)}{\tau}\right) h_0(x; C) dx - c + \int_0^{b^0} \alpha^e(C+1) h_0(x; C) dx \\
&\quad + \int_0^{b^0} \frac{1}{\tau} \left(\ln\left(\frac{2C+x}{2C+1}\right) + \frac{1}{2C+1}\right) h_0(x; C) dx, \alpha^e(C) \Bigg\}
\end{aligned}$$

where $\alpha^e(C) = \int \alpha(\epsilon, C) dF_\epsilon(\epsilon)$. It thus remains to be shown that b^0 is independent of $\bar{r}$.

III. Optimal Bidding Strategy.

Taking derivatives with respect to b^0 leads to the following FOC:

$$v + \epsilon + \frac{\ln\left(b^{0*}\right)}{\tau} - \frac{1}{\tau(2C+1)} + \ln\left(\frac{2C+1}{2C+b^{0*}}\right) \tau^{-1} - \alpha^e(C+1) = 0.$$

From here it can already be seen that b^{0*} is independent of $\bar{r}$. This completes the proof that the guess for the value function was correct.

Evaluating the FOC at $b^0 = e^{-\tau b}\bar{r}^{-1}$ and solving for b gives:

$$b^* = \frac{1}{\tau}\ln\left((2C+1)e^{-\frac{1}{2C+1}}e^{\tau(v+\epsilon-\alpha^e(C+1))} - 1\right) - \frac{\ln(2C\bar{r})}{\tau}.$$

Rearranging leads to equation (12). $\square$

References

Burdett, Kenneth and Tara Vishwanath, "Declining Reservation Wages and Learning," *Review of Economic Studies*, 1988, *55*, 655–666.

David, H. A. and H. N. Nagaraja, *Order Statistics*, Wiley-Interscience, 2003.

DeGroot, Morris H., *Optimal Statistical Decisions*, McGraw-Hill, 1970.

El-Gamal, Mahmoud A. and David M. Grether, "Are People Bayesian? Uncovering Behavioral Strategies," *Journal of the American Statistical Association*, September 1995, *90* (432), 1137–45.

Gumbel, Emil Julius, *Statistics of Extremes*, Columbia Univ. Press, 1958.

Kotz, Samuel and Saralees Nadarajah, *Extreme value distributions. Theory and applications*, London: Imperial College Press, 2002.

Sailer, Katharina, "Searching the eBay Marketplace.," *mimeo*, 2005, (August).

Weber, Robert J., "Multiple-Object Auctions," in Paul Klemperer, ed., *The economic theory of auctions*, VOL 2, Cheltenham, UK and Northampton, MA: Elgar, 2000, pp. 240–66.