

# Bayesian Estimation of Bid Sequences in Internet Auctions Using a Generalized Record Breaking Model

Eric T. Bradlow and Young-Hoon Park\*

December 2005

---

\*Eric T. Bradlow is The K.P. Chao Professor, Professor of Marketing and Statistics and Academic Director of the Wharton Small Business Development Center, The Wharton School, University of Pennsylvania, 3730 Walnut Street, 761 JMHH, Philadelphia, PA 19104-6371; phone: (215) 898-8255; fax: (215) 898-2534; email: ebradlow@wharton.upenn.edu. Young-Hoon Park is Assistant Professor of Marketing, Johnson Graduate School of Management, Cornell University, 330 Sage Hall, Ithaca, NY 14853-6201; phone: (607) 255-3217; fax: (607) 254-4590; email: yp34@cornell.edu. The authors would like to thank seminar participants at Columbia University, 2004 Joint Statistical Meetings, 2004 Marketing Science Conference, and Syracuse University for their comments.

# Bayesian Estimation of Bid Sequences in Internet Auctions Using a Generalized Record Breaking Model

## Abstract

The motivation for this research is derived from a currently active research topic in both marketing and economics, Internet auctions. A sequence of bids in this market mechanism can be viewed as record breaking events with participants (bidders) having dynamically changing valuations for the auctioned item, but the latent number of bidders “competing” in those events is unseen.

Our approach to addressing this problem is through data augmentation in which we stochastically draw the number of latent bidders, and then model the observed bid amounts and bid submission times conditional on the non-seen events.

In addition, as our research is among the first attempts to take a descriptive and exploratory modeling approach to auction data by allowing for dynamically changing (bid-by-bid) bidder valuations, we provide a model that is suggestive for those doing theoretical and/or theoretically grounded empirical work in this area.

**KEYWORDS:** Bidding dynamics, Data augmentation, Generalized record breaking, Latent bidders

# 1 Introduction

The growing importance of online auctions as exchange mechanisms is evident in the emerging auction literature from economists (e.g., Bajari and Hortaçsu 2003, Roth and Ockenfels 2002) to statisticians (Shmueli and Jank 2005), to behavioral researchers (e.g., Ariely and Simonson 2003, Greenleaf 2004, Kamins et al. 2004). We present here a new approach to modeling auction data (the entire sequence of bids and bid times) that relies on merging (yet also modifying) an existing record breaking literature in statistics (e.g., Carlin and Gelfand 1993) with Bayesian data augmentation methods (Albert and Chib 1993, Tanner and Wong 1987).

Record breaking data, in which only those data points that break the current record are observed (among those accumulated sequentially over time) have been well studied in a variety of practical situations such as World and Olympic sporting records (e.g., Carlin and Gelfand 1993, Robinson and Tawn 1995) and meteorological data (e.g., Brown and Katz 1995, Coles and Tawn 1996). Our link to this literature is to note that the arrival of a new bid for an auction item can be considered a “record breaking event”; that is, the current bid has broken the record of the last bid.

While this link is tight, the existing record breaking literature is deficient to examine auction data in a number of important ways. The literature has been concerned with independent events and a constant (stationary) mean (e.g., Tryfos and Blackmore 1985). In our application area, assuming stable bidder valuations, while commonly assumed in rational economic models (e.g., Guerre et al. 2000, Hendricks et al. 2003), has been questioned in a variety of empirical settings (e.g., Ariely and Simonson 2003, Chakravarti et al. 2002, Park and Bradlow 2005). Further, the literature has directly incorporated record breaking opportunities wherein occurrence times of all failed record breaking attempts are known (e.g., Carlin and Gelfand 1993). In auctions, however, the assumption of perfect knowledge about the “failures” could be very restrictive, which represents the unique merging of the record breaking and the data augmentation literature developed here.

We apply the proposed modeling framework of records with unseen intermittent events to

auction data in which bid amounts and bid times are analogous to record breaking events and the event times. Bid amounts (and bid times) are accumulated sequentially, and are recorded as long as a bid amount exceeds the outstanding bid. That is, (observed) bid increments have to be greater than zero. Moreover, the number of latent competing bidders, i.e., record breaking attempts, is not observed (or there exists an unseen competition set as in Bradlow and Fader 2001). Instead, we only observe the largest of the potential bids, but not those by the latent bidders. We therefore investigate stochastic versions of the classical record breaking problem for which we apply Bayesian estimation to predict observed bids and bid times. In this respect, our research appears to be a clear generalization of the research on record breaking events. To the best of our knowledge, it is the first time the record breaking paradigm has been applied to auction data.

We note that our stochastic modeling approach considerably differs from the empirical auction literature in economics. Specifically, our research is an attempt to add to the growing body of literature by providing a descriptive and exploratory look at auction data utilizing a flexible and parsimonious stochastic model. We note that our model does not rely on theoretical assumptions of utility maximization or strategic behaviors that have provided a significant path forward in this research area (e.g., Jofre-Bonet and Pesendorfer 2003, Pakes et al. 2005). Instead, our work is meant to extend empirical understanding of bidding behavior and provide a platform by which structural aspects can be incorporated into a stochastic approach. In this vein, the closest paper to ours is Park and Bradlow (2005), however their work assumes that the competing set of bidders is known based upon similar auctions that are on sale concurrently. In contrast, we stochastically impute the latent competition set, decompose those bidders into new entrants at a given bid and remaining bidders from the previous bid, and derive the distribution of the largest order statistic (Johnson et al. 1994, 1995), conditionally on having imputed the unseen attempts.

The remainder of the paper is organized as follows. Section 2 gives an overview of the data and describes exploratory analyses. In section 3, we provide a detailed specification of our Bayesian

model. Section 4 has model results and inferences. Section 5 concludes and suggests future research directions.

## 2 Data Description

We obtained a database of notebook computer auctions from one of the largest Internet auction sites in Korea. The auction mechanism on this site uses an ascending first-price or English auction in which the highest bidder wins and pays the amount s/he bids, and is essentially identical to the format used by Yahoo!Auctions. The database contains information regarding the complete history of bids, auction design features set by the seller, seller characteristics and product specifications. The total number of notebook auctions considered here is 218 with 3124 bids. Thus, on average, there are about 14.33 bids (std. dev. = 11.22), i.e., record breaking events, per auction.

There are three auction design variables under the sellers' control: placement (yes or no) of product images on the listing page, minimum bid ("public" reserve price) and auction duration. These variables are used as descriptors and are treated as exogenous.<sup>1</sup> In addition, we utilize seller reputation ratings given by past successful bidders. The rating is in the form of a positive, negative or neutral response after each transaction which we operationalize by using  $\log(\text{positive ratings} + 1)$  and  $\log(\text{negative and neutral ratings} + 1)$ .

The product specifications contain the following variables: (1) CPU type (Pentium or Celeron), (2) CPU speed, (3) memory, (4) hard disk, (5) screen size, (6) the number of months that the auction item has been used, and (7) brand name. There are two American (Compaq and IBM), two Japanese (Fujitsu and Sony) and two Korean brands (Sambo and Samsung), which account for about 31%, 11% and 52% of the 218 items, respectively. All of the rest of the brands, were aggregated and grouped into a category "others". Table 1 reports detailed descriptions of auction design, seller reputation ratings and product specifications. These variables, along with bid-specific

---

<sup>1</sup>Current research in both economics and marketing has focused on the so-called endogeneity problem both from an instrumental variables perspective (e.g., Vilcassim et al. 1999, Villas-Boas and Winer 1999) and from a modeling perspective (e.g., Manchanda et al. 2004) which suggests an excellent area to extend our modeling framework.

(time-varying) variables to be described, will serve as covariates towards explaining the latent number of competing bidders, the magnitude of new record breaking events (bid amounts), the bid times, and the variance (dispersion) of the underlying dynamic bidder valuations.

**Insert Table 1 about here**

In Internet auctions, incremental bid amounts are of great importance since the key decision by potential bidders centers on how much more to bid. Thus, our primary dependent variable is the bid increment, and specifically,  $\log(\text{bid increment})$ . Figure 1 shows a histogram of  $\log(\text{bid increment})$ , measured in Korean currency (won), where a fairly regular pattern of bid amounts<sup>2</sup> is observed.

**Insert Figure 1 about here**

Besides the amount of bids, bid submission times (or bid time increments) are of major interest both for sellers and potential bidders. Figure 2 shows that bidding activity is highly time concentrated primarily near the end of each auction: About 40% (50%) of bids are submitted within a one (two) hour window. This practice of a last-minute bidding phenomenon has attracted a good deal of attention among academic researchers (e.g., Bajari and Hortacısu 2003, Roth and Ockenfels 2002). Thus, we also model the bid time; hence a joint model of the amount and timing of new record breaking events.

**Insert Figure 2 about here**

Next, we build an explicit parametric and descriptive model of bid amount and time increments by inferring the number of latent potential bidders between subsequent “records”. Along this line, we investigate the determinants of increments in bid amounts and bid times through various auction-specific and bid-specific covariates.

---

<sup>2</sup>\$1 corresponded approximately to 1200 won at the time this data was collected.

### 3 A Bayesian Record Breaking Model

We first present a descriptive interpretation of our model and its assumptions. In section 3.2, we describe the heart of our Bayesian model, a time-varying latent process. A description of the priors and computational details are provided in sections 3.3 and 3.4 respectively. A simulation study to demonstrate the efficacy of our approach is given in section 3.5.

#### 3.1 Model Description

At the heart of our model is a latent record breaking process where we observe only the maximum amount and its bid time. We assume that there are a maximum of  $N_i$  latent bids (or bidders) for auction  $i$ .<sup>3</sup> For auction  $i$  and bid  $k$ , two types of latent bidders exist: a mixture of  $n_{i,k}$  *new* entrants, and  $n_{i,k}^*$  *remaining* bidders who carry-over from prior to bid  $k$  (the process is described below). Let  $\bar{n}_{i,k} = n_{i,k} + n_{i,k}^*$  denote the *total* number of latent bidders. Thus, our model considers bidding participation behavior wherein latent bidders come in ( $n_{i,k}$  of them), stay ( $n_{i,k}^*$ ) or go out of the bidding process. In particular, our stochastic auction model proceeds as follows.

1. Imagine that the  $(k - 1)$ -st bid for auction  $i$  has just arrived at time  $t_{i,k-1}$  and that there is  $(T_i - t_{i,k-1})$  time left in the auction where  $T_i$  is the known and set auction duration.
2. A latent competition set (maximum size  $N_i$  as above) of asymmetric bidders (a mixture of new entrants and remaining bidders as above) determine (in isolation) their expected valuations for auction  $i$  starting at time  $(t_{i,k-1} + \varepsilon)$  until the next bid (if it does) arrives. We denote the observed bidder as the  $(\bar{n}_{i,k} + 1)$ -st one.
3. Each of the  $n_{i,k}$  new latent bidders have bidding intensities given by  $\mu_{i,k}$  for bid increments through a truncated log-normal distribution, and by  $\tau_{i,k}$  for bid time increments through the truncated exponential distribution. For the  $n_{i,k}^*$  remaining bidders, their bidding intensities

---

<sup>3</sup>We examined that our model is not particularly sensitive to the choice of maximal  $N_i$  used for computation. Furthermore, it can be made arbitrarily large.  $N_i$  is included to allow for the possibility of a known competition set. We thank an anonymous reviewer for a comment that led to this choice.

are given by  $\mu_{i,k,j}^*$  for bid increments through a truncated log-normal distribution and  $\tau_{i,k}^*$  for bid time increments.

4. Let  $t_{i,k} \in (t_{i,k-1}, T_i)$  denote the arrival time of bid  $k$  and  $b_{i,k, [\bar{n}_{i,k}+1]}$  the amount of that bid.

That is, our model proposes (as given in section 3.2) that the observed bid time increment is the one among the latent bidders with the smallest “arrival” time which is determined by the minimum of  $(\bar{n}_{i,k} + 1)$  exponential bidders and the value of that bid increment is determined by  $\mu_{i,k}$  and  $\mu_{i,k,j}^*$ . We next provide a formal layout of the model.

### 3.2 Model Development

More formally, we model the number of latent entrant bidders,  $n_{i,k}$ , who enter auction  $i = 1, 2, \dots, I$  after bid  $(k - 1)$ ,  $k = 1, 2, \dots, K_i$ ,  $n_{i,k}$ , by a truncated Poisson distribution with mean  $\lambda_{i,k}$ :

$$f(n_{i,k} | \lambda_{i,k}) = \frac{\lambda_{i,k}^{n_{i,k}} \exp(-\lambda_{i,k})}{(n_{i,k})!}, \quad n_{i,k} \in [0, N_i - n_{i,k}^*]. \quad (1)$$

To model the time-varying mean (within auction and across bids),  $\lambda_{i,k}$ , we choose an identical regression specification (same predictors yet varying coefficients) for all four core “kernels” of our model (described below): the distribution for the number of latent bidders, the distribution of bidder valuations for bid increments and their variance, and the distribution for bid time increments. This unified approach allows us to assess the impact of variables on all key aspects simultaneously. More specifically, we model

$$\begin{aligned} \log(\lambda_{i,k}) &= X'_{i,k} \cdot \alpha + \epsilon_{i,k} \\ &= \alpha_0 + \alpha'_1 \text{AD}_i + \alpha'_2 \text{SR}_i + \alpha'_3 \text{PS}_i + \alpha'_4 \text{MC}_i + \alpha'_5 \text{BC}_{i,k} + \epsilon_{i,k}, \end{aligned} \quad (2)$$

where  $\text{AD}_i$  is a vector of design variables, i.e., placement of product images,  $\log(\text{minimum bid})$  and auction duration;  $\text{SR}_i$  is a vector of variables for seller reputation, i.e.,  $\log(\text{positive ratings} + 1)$  and  $\log(\text{negative and neutral ratings} + 1)$ ;  $\text{PS}_i$  is a vector of variables for product specifications, i.e., CPU type, CPU speed, memory, hard disk, screen size, the number of months for use and

brand names as in Table 1;  $MC_i$  is a vector indicating the degree of marketing competition, i.e., breadth (how many items similar to the focal item are on sale concurrently) and depth (how many of those similar items also of the same brand as the focal item are on sale concurrently)<sup>4</sup>;  $BC_{i,k}$  is a vector of time-varying variables, i.e., number of bids before the  $k$ -th bid (i.e.,  $k - 1$ ), the previous incremental bid amount and bid time, and the amount of remaining time ( $T_i - t_{i,k-1}$ ) in the auction; and  $\epsilon_{i,k} \sim N(0, \sigma_\alpha^2)$ .

Given  $n_{i,k}$  latent entrant bidders as in equations (1) and (2) and  $n_{i,k}^*$  latent bidders who remain from the  $(k - 1)$ -st bid, in the participation process at bid  $k$  for auction  $i$ , this implies that there are  $\bar{n}_{i,k}$  unseen bids which we denote  $b_{i,k,j}$ ,  $j = 1, \dots, \bar{n}_{i,k}$ . Define  $b_{i,k,[j]}$  as the  $j$ -th corresponding bid order statistic at bid  $k$  for auction  $i$ . Therefore,  $b_{i,k,[\bar{n}_{i,k}+1]}$  is the  $k$ -th observed bid and  $b_{i,0,[n_{i,0}+1]}$  is the minimum opening bid in auction  $i$ . We define random variables  $y_{i,k,j} = b_{i,k,j} - b_{i,k-1,[\bar{n}_{i,k-1}+1]}$  as the latent bid increments, and  $y_{i,k,[\bar{n}_{i,k}+1]}$  the observed bid increment at bid  $k$  for auction  $i$ .

To complete the model description, we need to describe the process by which latent bidders are assumed (if at all) to “carry-over” and remain after a bid occurs. While there were numerous ways to do this, we believe our choice is both plausible while at the same time an area for future research. In particular, we use a latent threshold model for continued participation. That is, let  $z_{i,k,j}$  be 1 if the  $j$ -th latent bidder from the  $(k - 1)$ -st bid remains in bid  $k$  for auction  $i$ , and 0 otherwise. We operationalize  $z_{i,k,j}$  to be 1 if  $y_{i,k,j} \geq \theta \cdot y_{i,k,[\bar{n}_{i,k}+1]}$ ,  $0 \leq \theta \leq 1$ . That is,  $n_{i,k}^* = \sum_{j=1}^{\bar{n}_{i,k}-1} z_{i,k,j}$  latent bidders who are “ $\theta$ -close” competitors of the observed bidder remain. As a result of this component, the latent bidders after each bid are no longer symmetric, nor are they all assumed to be new. We describe below how these bidder types are modeled with different parameters.

The proposed model assumes these bid increments have a log-normal distribution truncated

---

<sup>4</sup>We followed Chan et al. (2005) who operationalized these measures via auction-specific variables. First, we pooled all items being auctioned during the time interval that the focal item is being auctioned. We calculate the mean CPU speed, memory, hard drive, and screen size of these items. We next select items that are within one standard deviation of the focal item’s speed, memory, hard drive, and screen size. We then match the items on the discrete attributes. We define these items as comprising the “similar item superset.”

below by zero (bids are ascending) with density functions, for  $j = 1, \dots, (\bar{n}_{i,k} + 1)$ ,

$$f(y_{i,k,j} | \mu_{i,k}, \sigma_{i,k}^2) = \frac{1}{[2\pi\sigma_{i,k}^2]^{1/2}} \exp \left\{ -\frac{1}{2\sigma_{i,k}^2} (y_{i,k,j} - \mu_{i,k})^2 \right\} / [1 - \Phi(\frac{\mu_{i,k}}{\sigma_{i,k}})] . \quad (3)$$

The ability of our model to describe and predict the bid increments (record breaking events) is thus due to our ability to flexibly specify a model for  $\mu_{i,k}$ , the mean of the bid increments at time  $(t_{i,k-1} + \varepsilon)$ .

Specifically, we model the time-varying (bid  $k$ ) mean vector for the  $(n_{i,k} + 1)$  latent bidders as:

$$\mu_{i,k} = X'_{i,k} \cdot \beta + \varsigma_{i,k} , \quad \varsigma_{i,k} \sim N(0, \sigma_{\beta}^2) , \quad (4)$$

where  $\beta = (\beta'_0, \beta'_1, \dots, \beta'_5)$ ; an identical regression specification as in equation (2), albeit with different coefficients, with the exception of the inclusion of a random error term  $\varsigma_{i,k}$  allowing for the possibility of omitted variables.<sup>5</sup> For the  $n_{i,k}^*$  carry-over latent bidders, we modify the mean valuation vector in equation (4) to account for the fact that the mean bid increment for remaining bidders may be affected by their previous “latent” bid increments. We thus build in an AR(1) process where we allow for the possibility that previous latent bidders “anchor” in some way on their previous valuations:<sup>6</sup>

$$\mu_{i,k,j}^* = X'_{i,k} \cdot \beta + \beta_6 \cdot y_{i,k-1,j} \cdot z_{i,k,j} + \varsigma_{i,k} , \quad \varsigma_{i,k} \sim N(0, \sigma_{\beta}^2) . \quad (5)$$

Note that once augmented with latent bid increments  $y_{i,k,1}, \dots, y_{i,k,\bar{n}_{i,k}}$ , the conditional distribution of the largest bid increment, i.e., the observed bid  $y_{i,k, [\bar{n}_{i,k} + 1]}$ , given the unseen bid increments, is given by a truncated log-normal distribution truncated below at the largest of the unseen bid increments. Furthermore, the conditional distribution of the unseen bid increments, conditional on the observed bid increment, is also truncated log-normal, now truncated below by zero and above by the observed bid increment; hence the computational simplicity of our augmented data model.

---

<sup>5</sup>We fit our model both with and without this error term on all parts of the model. For bid increments and bid time increments, it improved the fit of the model and hence was included.

<sup>6</sup>We assume the same error term for  $\mu_{i,k}$  and  $\mu_{i,k,j}^*$ , for model parsimony.

We also posit an identical regression specification for the auction by bid-specific variances,  $\sigma_{i,k}^2$ , from equation (3) given by:

$$\log(\sigma_{i,k}^2) \sim N(X'_{i,k} \cdot \zeta, \sigma_\zeta^2), \quad (6)$$

a log-normal regression. Again, such a specification allows us to flexibly model the uncertainty in bid increments using the same set of covariates.

Finally, we now uncondition on the arrival time of the  $k$ -th bid,  $t_{i,k}$ , by building a predictive model for its arrival (if it occurs). As described in section 3.1, if we have  $(\bar{n}_{i,k} + 1)$  potential bidders, each with truncated exponential arrival “speeds”, the observed bid time is the value of the minimum order statistic, “the fastest participant”.<sup>7</sup> That is, the smallest of the  $t_{i,k}$ , the observed bid time after the  $(k - 1)$ -st bid for auction  $i$ , is given by:

$$\Pr[\min(t_{i,k,j}) - t_{i,k-1} = \Delta t_{i,k}] = \frac{[(\bar{n}_{i,k} + 1) \cdot \tau_{i,k}] \times \exp \left\{ - [(\bar{n}_{i,k} + 1) \cdot \tau_{i,k}] \cdot \Delta t_{i,k} \right\}}{1 - \exp \left\{ - [(\bar{n}_{i,k} + 1) \cdot \tau_{i,k}] \cdot (T_i - t_{i,k-1}) \right\}}, \quad (7)$$

where

$$\log(\tau_{i,k}) = X'_{i,k} \cdot \gamma + \xi_{i,k}, \quad \xi_{i,k} \sim N(0, \sigma_\gamma^2). \quad (8)$$

As shown in equation (7), the probability of when bid  $k$  occurs is also truncated exponential with rate equal to the sum of  $\tau_{i,k}$  across all the latent and observed bidders.<sup>8</sup> This result, that the minimum of a set of exponentials is exponentially distributed with rate equal to the sum of the rates is a nice property of the exponential distribution. It is important to note that our choice of the exponential distribution comes with a balance of positive and negative aspects. In particular, the positive aspects are that it leads to closed-form solutions to the model, a feat that should not be minimized. On the other hand, a more general timing model like the Weibull model would be more flexible; albeit, at the cost of closed-form tractability. However, the error term in (8) affords us a great degree of flexibility too.

---

<sup>7</sup>Our model can be thought of as a continuous time model where bids arrive according to a continuous exponential process with discrete “restarts” at the observed bids. We approximate that the sum of the rates for each time period starts  $\varepsilon$  after the previous bid, i.e., bidders immediately update.

<sup>8</sup>We only included a parameter corresponding to the magnitude shift of the AR(1) process for bid increments. We note that the AR(1) process could also be included for bid time and variance which we leave for future research.

### 3.3 Prior specification

All slope parameters which govern the time-varying number of latent bidders,  $\alpha$ , mean log(bid increment),  $\beta$ , bid increment variances,  $\zeta$ , and mean log(rate of bid speed),  $\gamma$  are given slightly informative but vague priors,  $N(0, \text{precision} = 0.00001)$ , to ensure proper posteriors but also to allow the data to primarily govern the inferences. Sensitivity analyses by varying these precisions indicated no impact due to the large sample sizes that are being used here. All variance components,  $\sigma_\beta^2$ ,  $\sigma_\zeta^2$ , and  $\sigma_\gamma^2$  are given slightly informative inverse-gamma priors (Gelman et al. 1995), with shape and scale set at 0.01.

### 3.4 Computational Approach

Inferences were obtained using a data augmentation MCMC sampler, implemented in the freely available software Bayesian Inference Using Gibbs Sampling, WinBUGS. Results reported are from the output of three independent chains run for 10000 iterations each started from hyper-dispersed starting values, with a burn-in period of 7500 iterations and utilizing the 7500 draws (2500 per chain) thereafter. Convergence was diagnosed both graphically and using the  $F$ -statistic diagnostic of Gelman and Rubin (1992). We believe it is of importance to note how the model can be fit in WinBUGS (with ease) once the data augmentation scheme was utilized. To enable use by other academicians and practitioners, the WinBUGS code will be posted to an accessible website.

### 3.5 Simulation Experiment

Due to the complex nature of the model, and the latent processes that are specified within, we conducted a small scale MCMC simulation experiment to understand the identifiability of our proposed model. To accomplish this, we simulated 10 (replicate) data sets using a reduced set of covariates (described below), but fully parameterized and structured version of our model. The values for the simulation parameters (intercepts and slopes) were chosen to mimic the values found in our real data. Other simulation values were chosen to mimic our data set in size and nature.

Further simulations not reported suggest that parameter recoverability is robust, at least for the other limited conditions that we tested.

In particular, the ten simulated data sets were generated as follows. Each consisted of data for 200 auctions and 15 bids per auction. For each of the four model components, we utilized an intercept, one static (auction-specific) covariate, and one time-varying (bid-specific) covariate. This structure was chosen as it represents the basic structure of each aspect of our latent linear regressions: an intercept, static covariates, and time-varying covariates. In addition, for the latent bid increments we also included another parameter corresponding to the magnitude shift of the AR(1) process to determine those bidders who remain from the previous bid and a corresponding threshold parameter ( $\theta$ ). Thus, this simulation structure is fully general.

The results reported in Table 2 are the average over the 10 simulations.<sup>9</sup> As shown in Table 2, all of the parameters are contained in their respective average 95% posterior intervals (and were for all 10 simulations) suggesting the clear identifiability of the model.

**Insert Table 2 about here**

## 4 Model Results

Before presenting detailed results for our fully general model, we demonstrate its superiority in comparison to a number of interesting nested models, as well as a simple benchmark model. In particular, our model in section 3 considers bidding participation wherein latent bidders come in, stay or go out of the bidding process. This model, however, nests a model where all bidders once they arrive stay throughout, for subsequent periods as operationalized here ( $\theta = 0$ ) as well as a model where latent bidders are independent across bids, i.e., a new set of latent bidders arrives at each bid ( $\theta = 1$ ). We also performed a series of OLS regression analyses of bid and time increments as a function of the same covariates. Table 3 reports that our general model performs best based on the mean absolute prediction error (MAPE) in bid increments as well as time increments. Given

---

<sup>9</sup>Standard errors for each result are available from the authors upon request.

the theoretical completeness as well as its model fit, we focus on results of our general model (Model 3 in Table 3) hereafter.

**Insert Table 3 about here**

There are four major areas that we report upon in summarizing our results. First, we describe inferences for the latent number of competing bidders (both new and remaining), as inferred by the posterior distribution of the  $n_{i,k}$  and  $n_{i,k}^*$ . These are particularly informative as it represents a unique contribution of our model formulation. Second, we describe inferences relating  $n_{i,k}$  to the covariates  $X_{i,k}$ , i.e.,  $\alpha$  as in equation (2). Third, we describe inferences that can be obtained by looking at summaries (posterior means) of the posterior distribution for  $\beta$  and  $\gamma$ , the parameters governing the distribution for bid and time increments, respectively, and for  $\zeta$ , the parameters governing the distribution of bidder valuation variances. Finally, we report findings on assessing the fit of the model by looking at the posterior predictive distribution (Gelman et al. 1996) for a variety of different quantiles along the observed empirical distribution of bid and time increments.

#### 4.1 Latent Bidders

Besides computational convenience due to closed-form conditional distributions, understanding the latent bid and time increments provides a unique look into the underlying latent bidder process and hence the auction as a whole. To summarize the inferences for the number of latent bidders (both new and remaining), we describe summaries of the posterior distributions of the  $n_{i,k}$  and  $n_{i,k}^*$  themselves, in particular their posterior means denoted  $\tilde{n}_{i,k}$  and  $\tilde{n}_{i,k}^*$ , and posterior means of the total number of latent bidders,  $\bar{n}_{i,k}$ , as a fraction of total bids,  $\frac{\max(\tilde{n}_{i,k}, \forall k)}{\max(\tilde{n}_{i,k}, \forall k) + K_i}$ , where  $K_i$  is the number of observed bids in auction  $i$ . This auction-level statistic provides an ‘‘impact’’ measure on the role (fraction) of total latent to observed bidders (bids) in the auction process across auctions.<sup>10</sup>

---

<sup>10</sup>Note that in our model, we do not differentiate between bids and bidders as our interest is in the bid increments. If one is interested in the bidders and their increments, one might look at  $\frac{\max(\tilde{n}_{i,k}, \forall k)}{\max(\tilde{n}_{i,k}, \forall k) + U_i}$  instead, where  $U_i$  is the number of unique bidders.

An interesting and unique aspect of findings surrounds the distribution of the  $\tilde{n}_{i,k}$  and  $\tilde{n}_{i,k}^*$ , the posterior means of the number of latent bidders (new and remaining) across the 3124 bids in 218 auctions. In Table 4, we present a summary of the posterior means and 95% posterior intervals for  $n_{i,k}$  and  $n_{i,k}^*$  corresponding to the minimum, 2.5%-ile, 25%-ile, median, 75%-ile, 97.5%-ile and maximum values of  $\tilde{n}_{i,k}$  and  $\tilde{n}_{i,k}^*$ . We see that the posterior mean of the latent number of new (remaining) bidders  $\tilde{n}_{i,k}$  ( $\tilde{n}_{i,k}^*$ ) ranges from 0.38 (0.15) to 33.17 (53.93) indicating significant variation in those values.

**Insert Table 4 about here**

An analysis of the latent bidders, by looking at the posterior means of the ratio of the total number of latent bidders to the total number of bids (seen= $K_i$  and unseen= $\tilde{n}_{i,k}$ ),  $\frac{\max(\tilde{n}_{i,k}, \forall k)}{\max(\tilde{n}_{i,k}, \forall k) + K_i}$ , is presented in Figure 3. It shows that the latent number of bidders play a significant role, i.e., the latent bidders are significant in number relative to the observed bidders. Also, we note that the correlation between the predicted number of latent bidders and number of actual bidders across auctions is 0.563, a finding suggesting that while auctions with larger observed bidders tend to have more latent bidders, there is still significant variation. Furthermore, we provide in Figure 4 the posterior means of the ratio of the latent number of remaining bidders to the number of total latent bidders (both new and remaining). As shown in Figure 4, the latent number of remaining bidders is considerably smaller than the latent number of new entrants (on average), but is still an important consideration.

**Insert Figures 3 and 4 about here**

Figures 5 and 6 show a more detailed analysis of the bivariate relationships between the latent number of remaining bidders and observed log(bid increment) and time increment, respectively. In each figure, we break the observed x-axis measure (log (bid increment) or bid time increment) into deciles and then do a boxplot of the posterior means of the latent number of remaining bidders for each decile. Both figures show a clear downward pattern in  $\tilde{n}_{i,k}^*$  over deciles which implies

that larger bid and time increments significantly influence the bidding participation behavior of the remaining bidders. This is corroborated by a very high threshold estimate (posterior mean of  $\theta = 0.959$ ) in our estimation.

**Insert Figures 5 and 6 about here**

Another set of inferences related to the latent number of bidders, in this case new entrant bidders  $n_{i,k}$ , of particular interest is how the number of latent entrant bidders varies with the covariates as given in Table 1 and equation (4). The results are presented in the first column in Table 5 and are derived from the posterior means  $\tilde{\alpha}$  which describe  $\log(\lambda_{i,k})$ . First, the auction design variables,  $AD_i$ , are important descriptors in explaining the number of latent entrant bidders,  $n_{i,k}$ . They have significant impact, positive for image and negative for minimum bid amount and auction duration. That is, auction items with product images attract more latent bidders, and those that have a high “buy-in” (minimum bid), or a long wait to determine a winner (long auction duration), have lower values. Second, seller reputation ratings,  $SR_i$ , are significant predictors for bidder entry behavior. Specifically, the extent (number) of positive ratings is more important than that of negative ratings, and in the expected direction. Third, most variables related to the product specifications,  $PS_i$  are important predictors of entry,  $n_{i,k}$ , in a positive manner, except for usage which is negative. This is somewhat confirmatory, as one might expect that higher levels of product specification would attract greater attraction; albeit, depending on price, and that a good that is older (higher usage) would attract less.

**Insert Table 5 about here**

Finally, one important set of inferences centers on the bid-specific variables  $BC_{i,k}$ , i.e., the dynamics as the auction progresses: As the number of observed bids (i.e.,  $k - 1$ ) increases, there tends to be a greater number of latent entrant bidders. We also observe negative effects of the bid and time increments in the previous round and the remaining time in the auction duration. This suggests that when bid activity is clustered either in terms of bid amount and bid time, the

number of latent bidders increases. In sum, all time-varying variables are important predictors of the number of latent bidders and demonstrates the importance of the time-varying nature in our stochastic model.

## 4.2 Parameter Inferences

In addition to understanding the latent bidding process, one major goal is to understand the drivers of the observed bid increments, the source of variations in bidder valuations, and bid time increments. Such inferences are available by looking at summaries of elements of the posterior distribution of  $\beta$ ,  $\zeta$ , and  $\gamma$  respectively, as shown in Table 5.

The major findings suggested are as follows. Considering the three auction design variables,  $AD_i$ , we note a significant positive effect for having a product image on the listing page, indicating that our observed bid increments increase when potential bidders see the product they are going to bid on. This suggests that adding a product image may indeed raise confidence and perceptions of products, and as further corroboration, we also found lower bid increment variances ( $\zeta$  for presence of image). However, it did not affect the speed (bid time increments) at which bids arrive. Second, minimum bid is not a significant predictor for  $\mu_{i,k}$  or  $\log(\sigma_{i,k}^2)$  indicating a lack of effect on bid increments. However, it is negatively related to bid time increments, hence a larger minimum bid leads to smaller time increments, and thus a greater concentration of bids. Third, auctions of shorter duration have larger bid increments and marginally larger variations. We note that without explicit modeling equilibrium-generating process on the part of bidders, one should be cautious about interpreting these findings for “optimal” auction design; rather, as stated earlier, we consider these useful for further theoretical and/or empirical study.

Among variables of seller reputation,  $SR_i$ , we find that the amount of negative plus neutral ratings received from past bidders has a very significant negative effect on the bid increment and that positive seller reputation has a very significant positive effect on bidding behavior; yet no impact on the speed of the bid increments. The magnitude of positive reputation is much larger

than that of the negative reputation. Furthermore, both positive and negative seller reputation ratings reduce variance in bidding behavior, suggesting that they are considered as viable sources of information (variance reduction) by the bidders. We note that ours is the first research to simultaneously look at the impact of reputation on attracting bidders ( $\tilde{\alpha}$ ), bid increments ( $\tilde{\beta}$ ), bidder uncertainty ( $\tilde{\zeta}$ ), and bid time increments ( $\tilde{\gamma}$ ).

Third, we note very significant positive effects for CPU type (Pentium having larger bid increments than Celeron) and CPU speed in mhz on bid increments and bid time increments, which given their importance in defining notebook computers, is expected. That is, more valuable items tend to have larger bid increments and larger gaps between those bids. We note that this story is played out through all of our parameter estimates as  $\tilde{\alpha}$  is negative,  $\tilde{\beta}$  is negative,  $\tilde{\zeta}$  is positive and  $\tilde{\gamma}$  is positive for  $y_{i,k-1,[n_{i,k-1}+1]}$ . We also observe a significant positive effect on bid increments for the size of the hard disk (in gigabytes), an expected effect of  $\exp(0.024) = 1.020$  for every 1 gigabyte increase in hard disk space, and similarly an increase in the bid time increments. We also observe a significant positive effect for screen size both for bid increments and variances, which is somewhat intuitive given the sizes considered here. One other thing of note is that these findings, in conjunction with those for the number of latent bidders, suggests more (less) attractive auction items attract more (less) bidders who might bid more (less). Certainly this set of patterns is apparent in this data; an area for future research would be to confirm this empirically in a broader setting.

Fourth, we observe significant negative effects for bid increments and variances in both market competition measures ( $MC_i$ ). This informs that both measures (breadth and depth) help bidders better determine bid amounts which in turn reduces the uncertainty in bid increments. This explanation is consistent with literature in psychology and marketing in terms of consumer consideration and choice set formation and decision-making. Fifth, we observe a positive effect on the size of bid increments for the (number of bids  $- 1$ ), i.e., effectively bid  $k$ , and a negative effect on bid

increments for remaining time; these are an identical pattern to what we found for latent bidders as well. We also find a significant positive effect for the AR(1) process which indicates that bid increments tend to be higher for continued participants, possibly due to their having not been the “record breaker” at the previous round.

Another set of findings of particular interest is to understand a brand equity (e.g., Elrod and Keane 1995). In our context, we do this in Figure 7 by looking at a brand map constructed using the brand intercept’s posterior means for the number of latent bidders and the size of bid and time increments for the corresponding brands. From this we see that Samsung, for example, increases both the latent number of bidders and the size of the bid increments suggesting “high” brand equity in online auctions. In fact, all of the named brands, as compared to the baseline “others” have higher bid increments; yet Samsung is the only one with increased latent bidders. We believe that this is the first time that brand maps have been applied in the Internet auction context and represents another contribution of our research.<sup>11</sup>

**Insert Figure 7 about here**

### 4.3 Model Fit

In order to assess our degree of confidence about the inferences derived under our model, we assess its fit using a variety of more detailed measures than MAPE given in Table 3.

As shown in Table 6, the model’s ability to recover observed  $\log(\text{bid increment})$  values at differing points along the distribution of observed  $\log(\text{bid increment})$  is excellent. The table is generated by comparing the observed  $\log(\text{bid increment})$  at the minimum, maximum, and five common percentiles, to posterior means from their respective posterior predictive distributions. This is crucial to assessing the quality of the model, as we are interested in providing a model that would allow for adequate predictions along the entire spectrum of bids.

---

<sup>11</sup>As shown in the figure, the brand intercepts for the number of latent bidders (ranging from -1.5 to 0.5) and bid increments (ranging from -1.0 to 3.0) are relatively larger than the brand intercepts for bid time increment (ranging from -1.0 to 0.5). This could be true because the bid timing is less of an issue in Internet auctions because bidders will be better off by bidding at the last minute of the auction duration.

**Insert Table 6 about here**

As shown Table 6, in all cases from the minimum to the maximum value of observed  $\log(\text{bid increment})$ , the model fits quite well, as the observed increment falls well within the posterior predictive interval. For the maximum bid increment, the model slightly under-predicts (i.e., the posterior mean is too low) the observed value. This suggests potentially an area for future model improvement in which we consider utilization of a longer-tailed distribution, as compared to log-normal for the bid increments, or potentially including a more extensive set of descriptors for the bid increments.

With regards to bid time increments, we produce the corresponding table (Table 7), demonstrating the models' ability to predict the bid times along its entire distribution. Again, looking at the results in Table 7 suggests that all observed bid time increments are well within their posterior predictive 95% intervals, albeit for longer bid time increments there is a slight underestimation.

**Insert Table 7 about here**

## 5 Conclusion

In this research, we have generalized the classic record breaking problem in which records in a sequence can contain an underlying trend, yet the number of latent participants (in our jargon bidders) and their arrival times are unobserved. The way in which we have done this, via data augmentation of the number of and the valuation of their bids, and arrival times, has led to closed-form conditional distributions for the observed and latent bidder valuations. One obvious benefit of this approach is computational simplicity, in this case the ability to run the model in the freely available software WinBUGS, and to avoid the common difficulty in dealing with the marginal distribution of the largest order statistic.

We note a number of caveats and limitations to our research. First, the data collected here are from ascending first-price auctions. This choice by the site in which the winner pays the

price that he or she bids, in contrast to second-price auctions in which the winner pays the second highest bidder's offer, has been under much discussion in the auction literature (e.g., Lucking-Reiley 1999). With that in mind, the empirical regularities found here are thought more of as suggestive than in any way definitive. Yet, we hope our model provides a framework for further empirical exploration. Second, the model developed here is descriptive and exploratory in nature and is not based on theoretical assumptions of auction behavior and equilibria in the extant economic auction literature. In particular, extant research suggests that bidders may act strategically, not choosing to reveal their true valuations when facing competition. The  $n_{i,k}$  and  $n_{i,k}^*$  latent bidders can be considered hypothetical competitors in the mind of a bidder to which he or she does not strategically respond to. However, our work provides a platform by which structural aspects can be incorporated into a stochastic approach. For instance, our threshold model provides one way to think about ongoing and future bidder behavior; but, in addition, models which incorporate forward-looking behavior could also be incorporated.

Third, our model only records information from the last bid for remaining latent bidders through the AR(1) process and assumes all others start anew. It would be desirable to track the entire bidding history of remaining latent bidders, albeit this would add complexity for uncertain gains. Fourth, we consider the standard auction formats here which represent the general structure of record breaking events, and do not address the buy-it-now feature of online auctions. While this would be very tangential to the main thrust of this paper, it could be an interesting future research area due to the growing importance of this auction mechanism. Lastly, the lack of explicit modeling of variables under the control of the seller as endogenous to his or her knowledge is an area for interesting extension in this research domain, and could be incorporated through instrumental variables, possibly latent (Ebbes et al. 2005), or through an extended likelihood (Manchanda et. al 2004).

## References

- Albert, James H. and Siddhartha Chib (1993), "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, 88 (422), 669-679.
- Ariely, Dan and Itamar Simonson (2003), "Buying, Bidding, Playing or Competing? Value Assessment and Decision Dynamics in Online Auctions," *Journal of Consumer Psychology*, 13 (1), 113-123.
- Bajari, Patrick and Ali Hortag su (2003), "The Winner's Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions," *RAND Journal of Economics*, 34 (2), 329-355.
- Bradlow, Eric T. and P. S. Fader (2001), "A Bayesian Lifetime Model for the "Hot 100" Billboard Songs," *Journal of the American Statistical Association*, 96 (454), 368-381.
- Brown, Barbara G. and Richard W. Katz (1995), "Regional Analysis of Temperature Extremes: Spatial Analog for Climate Change?," *Journal of Climate*, 8 (1), 108-119.
- Carlin, Bradley P. and Alan E. Gelfand (1993), "Parametric Likelihood Inference for Record Breaking Problems," *Biometrika*, 80 (3), 507-515.
- Chan, Tat Y., Vrinda Kadiyali, and Young-Hoon Park (2005), "Willingness to Pay and Competition in Online Auctions," Working Paper, Johnson Graduate School of Management, Cornell University.
- Chakravarti, Dipankar, Eric A. Greenleaf, Atanu Sinha, Amar Cheema, James C. Cox, Daniel Friedman, Teck H. Ho, R. Mark Isaac, Andrew A. Mitchell, Amnon Rapoport, Michael H. Rothkopf, Joydeep Srivastava, and Rami Zwick (2002), "Auctions: Research Opportunities in Marketing," *Marketing Letters*, 13 (3), 281-296.
- Coles, Stuart G. and Jonathan A. Tawn (1996), "A Bayesian Analysis of Extreme Rainfall Data," *Applied Statistics*, 45 (4), 463-478.
- Ebbes, Peter, and Michel Wedel, Ton Steerneman, and Ulf B ockenholt (2005), "New Evidence for the Effect of Education on Income: Solving Endogeneity with Latent Instrumental Variables," *Quantitative Marketing and Economics*, 3 (4), 365-392.,
- Elrod, Terry and Michael P. Keane (1995), "A Factor-Analytic Probit Model for Representing the Market Structure in Panel Data" *Journal of Marketing Research*, 32 (1), 1-16.
- Gelman, Andrew, John B. Carlin, Hal S. Stern and Donald B. Rubin (1995), *Bayesian Data Analysis*, London, John Wiley & Sons.
- Gelman, Andrew, Xiao-Li Meng and Hal S. Stern (1996), "Posterior Predictive Assessment of Model Fitness via Realized Discrepancies (with discussion)," *Statistica Sinica*, 6 (4), 733-807.
- Gelman, Andrew and Donald B. Rubin (1992), "Inference from Iterative Simulation Using Multiple Sequences," *Statistical Science*, 7 (4), 457-472.
- Greenleaf, Eric A. (2004), "Reserves, Regret, and Rejoicing in Open English Auctions," *Journal of Consumer Research*, 31 (2), 264-273.
- Guerre, Emmanuel, Isabelle Perrigne and Quang Vuong (2000), "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68 (3), 525-574.

- Hendricks, Kenneth., Joris Pinkse and Robert H. Porter (2003), "Empirical Implications of Equilibrium Bidding in First-Price Symmetric, Common Value Auctions," *Review of Economic Studies*, 70 (1), 115-145.
- Jofre-Bonet, Mireia and Martin Pesendorfer (2003), "Estimation of a Dynamic Auction Game," *Econometrica*, 71 (5), 1443-1489.
- Johnson, Norman L., Samuel Kotz and N. Balakrishnan (1994), *Continuous Univariate Distributions*, 2nd edition, Volume 1, New York, John Wiley & Sons.
- Johnson, Norman L., Samuel Kotz and N. Balakrishnan (1995), *Continuous Univariate Distributions*, 2nd edition, Volume 2, New York, John Wiley & Sons.
- Kamins, Michael A., Xavier Drèze, and Valerie S. Folkes (2004), "A Field Study of the Effects of Minimum and Reserve Prices on Internet Auction," *Journal of Consumer Research*, 30 (4), 622-628.
- Lucking-Reiley, David (1999), "Using Field Experiments to Test Equivalence Between Auction Formats: Magic on the Internet," *American Economic Review*, 89 (5), 1063-1080.
- Manchanda, Puneet, Peter E. Rossi and Pradeep K. Chintagunta (2004), "Response Modeling with Non-Random Marketing Mix Variables," *Journal of Marketing Research*, 41 (4), 467-478.
- Pakes, Ariel, Michael Ostrovsky and Steve Berry (2005), "Simple Estimators for the Parameters of Discrete Dynamic Games with Entry/Exit Examples," Working Paper, Department of Economics, Harvard University.
- Park, Young-Hoon and Eric T. Bradlow (2005), "An Integrated Model for Bidding Behavior in Internet Auctions: Whether, Who, When, and How Much," *Journal of Marketing Research*, 42 (4), 470-482.
- Robinson Michael E. and Jonathan A. Tawn (1995), "Statistics for Exceptional Athletics Records," *Applied Statistics*, 44 (4), 499-511.
- Roth, Alvin E. and Axel Ockenfels (2002), "Last-Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet," *American Economic Review*, 92 (4), 1093-1103.
- Shmueli, Galit and Wolfgang Jank (2005), "Visualizing Online Auctions," *Journal of Computational and Graphical Statistics*, 14 (2), 299-319.
- Tanner, Martin A. and Wing Hung Wong (1987), "The Calculation of Posterior Distributions by Data Augmentation (with discussion)," *Journal of the American Statistical Association*, 82 (398), 528-550.
- Tryfos, Peter and Russell Blackmore (1985), "Forecasting Records," *Journal of the American Statistical Association*, 80 (389), 46-50.
- Vilcassim, Naufel J., Vrinda Kadiyali and Pradeep K. Chintagunta (1999), "Investigating Dynamic Multifirm Market Interactions in Price and Advertising," *Management Science*, 45 (4), 499-518.
- Villas-Boas, J. Miguel and Russell S. Winer (1999), "Endogeneity in Brand Choice Models," *Management Science*, 45 (10), 1324-1338.

Variable	Mean	Std. Dev.
<u>Auction design</u>		
Image (Yes = 1, No = 0)	0.64	
log(Minimum bid)	10.48	3.42
Duration (days)	5.83	3.02
<u>Seller reputation</u>		
log(Positive reputations + 1)	0.46	0.81
log(Negative reputations + 1)	0.41	0.76
<u>Product specification</u>		
CPU (Pentium = 1, Celeron = 0)	0.81	
CPU speed (mhz)	525.74	211.04
Memory (megabytes)	111.09	57.62
Hard disk (gigabytes)	11.59	6.77
Screen size (inches)	12.93	1.32
Usage (months)	9.37	9.81
<u>Brand name</u>		
Compaq	0.16	
IBM	0.15	
Fujitsu	0.03	
Sony	0.08	
Sambo	0.11	
Samsung	0.41	
Others	0.06	

Note: Brand names are not included.

Table 1: Data Description

	Variable	True Value	Estimated Value	
			Post. Mean	95% Post. Interval
$\alpha$ in $\log(\lambda_{i,k})$	Intercept	3.00	2.78	[0.02, 4.41]
	Auction-specific	0.02	-0.02	[-0.20, 0.20]
	Bid-specific	0.02	0.07	[-0.35, 0.42]
$\beta$ in $\mu_{i,k}$	Intercept	3.00	3.04	[2.84, 3.24]
	Auction-specific	0.10	0.18	[-0.08, 0.35]
	Bid-specific	0.20	0.19	[0.18, 0.21]
	AR(1): $y_{i,k-1,j}$	0.50	0.48	[0.45, 0.51]
	$\theta$	0.50	0.51	[0.49, 0.53]
$\zeta$ in $\log(\sigma_{i,k}^2)$	Intercept	1.40	1.32	[1.00, 1.65]
	Auction-specific	0.40	0.30	[0.00, 0.60]
	Bid-specific	0.02	0.02	[0.00, 0.05]
$\gamma$ in $\log(\tau_{i,k})$	Intercept	1.20	1.37	[0.85, 1.92]
	Auction-specific	0.17	0.12	[-0.35, 0.54]
	Bid-specific	0.73	0.74	[0.70, 0.82]

Table 2: Simulation Results

	Regression	Model 1: $\theta = 0$	Model 2: $\theta = 1$	Model 3: $0 \leq \theta \leq 1$
log(Bid increment)	1.362	1.188	0.725	0.637
Bid time increment	0.399	0.347	0.345	0.299

Table 3: Model Fit: Mean Absolute Prediction Error

Quartile	$n_{i,k}$		$n_{i,k}^*$	
	Post. Mean	95% Post. Interval	Post. Mean	95% Post. Interval
Minimum	0.38	[0, 1]	0.15	[0, 1]
2.5%	0.68	[0, 2]	0.45	[0, 2]
25.0%	4.72	[1,12]	1.24	[0, 4]
50.0%	9.48	[2,24]	3.01	[0, 8]
75.0%	15.79	[3,41]	7.50	[2,14]
97.5%	25.60	[5,56]	48.27	[37,56]
Maximum	33.17	[6,59]	53.93	[48,59]

Table 4: Posterior Means and 95% Probability Intervals for Particular  $n_{i,k}$  and  $n_{i,k}^*$

Variable	$\tilde{\alpha}$ in $\log(\lambda_{i,k})$	$\tilde{\beta}$ in $\mu_{i,k}$	$\tilde{\zeta}$ in $\log(\sigma_{i,k}^2)$	$\tilde{\gamma}$ in $\log(\tau_{i,k})$
Intercept	2.965*	3.086*	1.417*	1.201*
AD <sub>i</sub>				
Image	0.134*	0.221*	-0.237*	0.124
log(Minimum bid)	-0.122*	0.047	-0.158	-0.243*
Duration	-0.151*	-0.109*	-0.175	-0.254
SR <sub>i</sub>				
log(Positive ratings +1)	0.086*	0.052*	-0.072*	0.066
log(Negative ratings + 1)	-0.047*	-0.030*	-0.128	-0.062
PS <sub>i</sub>				
CPU type	0.024	0.099*	0.418*	0.168*
CPU speed	0.093*	0.193	0.137*	0.048*
Memory	0.102*	0.039*	0.063*	0.020
Hard disk	0.107	0.020*	0.260	0.020*
Screen size	0.045*	0.026*	0.038*	0.072
Usage	-0.044	-0.073*	-0.063	0.025
MC <sub>i</sub>				
Breadth	-0.060	-0.025*	-0.165*	-0.005
Depth	-0.009	-0.032*	-0.068*	-0.051*
BC <sub>i,k</sub>				
$k - 1$	0.026*	0.018*	-0.021*	0.728*
$y_{i,k-1,[n_{i,k-1}+1]}$	-0.025*	-0.112*	0.035*	0.024*
$\Delta t_{i,k-1}$	-0.209*	0.027*	-0.088*	-0.071*
$T_i - t_{i,k-1}$	-0.559*	-0.225*	0.592*	0.405*
AR(1) <sub>i,k,j</sub>				
$y_{i,k-1,j}$		0.013*		

Notes: 1. Parameter estimates for brand names are not included.

2. \* indicates that zero lies outside of the 95% posterior interval.

Table 5: Posterior Means of Model Parameters

Quartile	Observed log(Bid increment)	Predicted log(Bid increment)	
		Post. Mean	95% Post. Interval
Minimum	5.704	5.700	[ 5.684, 5.713]
2.5%	6.908	6.905	[ 6.820, 6.952]
25.0%	8.517	8.521	[ 8.304, 8.638]
50.0%	8.556	8.648	[ 8.395, 8.971]
75.0%	10.309	10.250	[ 9.234,11.400]
97.5%	12.897	12.560	[ 9.716,15.480]
Maximum	14.604	13.560	[ 9.982,17.310]

Table 6: Observed log(Bid Increment) Versus Predicted log(Bid Increment)

Quartile	Observed time increment	Predicted time increment	
		Post. Mean	95% Post. Interval
Minimum	2.3E-05	0.0052	[ 0.0001, 0.0191]
2.5%	0.0003	0.0065	[ 0.0002, 0.0229]
25.0%	0.0107	0.0130	[ 0.0003, 0.0498]
50.0%	0.0787	0.0798	[ 0.0024, 0.2836]
75.0%	0.3405	0.3404	[ 0.0108, 1.0889]
97.5%	2.7150	1.7939	[ 0.0410, 4.8058]
Maximum	6.0978	3.7484	[ 0.1635, 8.7359]

Table 7: Observed Time Increment versus Predicted Time Increment

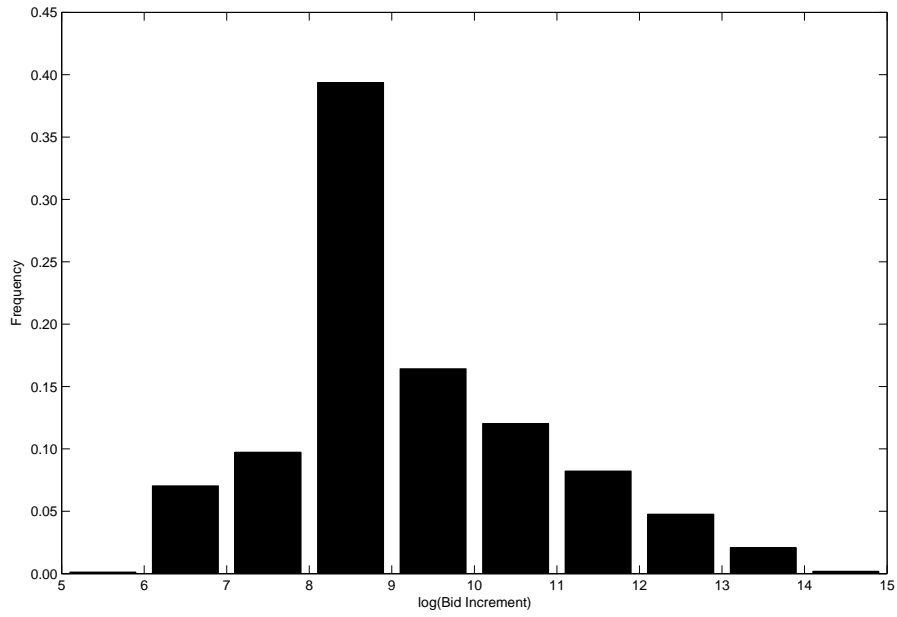


Figure 1: Histogram of  $\log(\text{Bid Increment})$

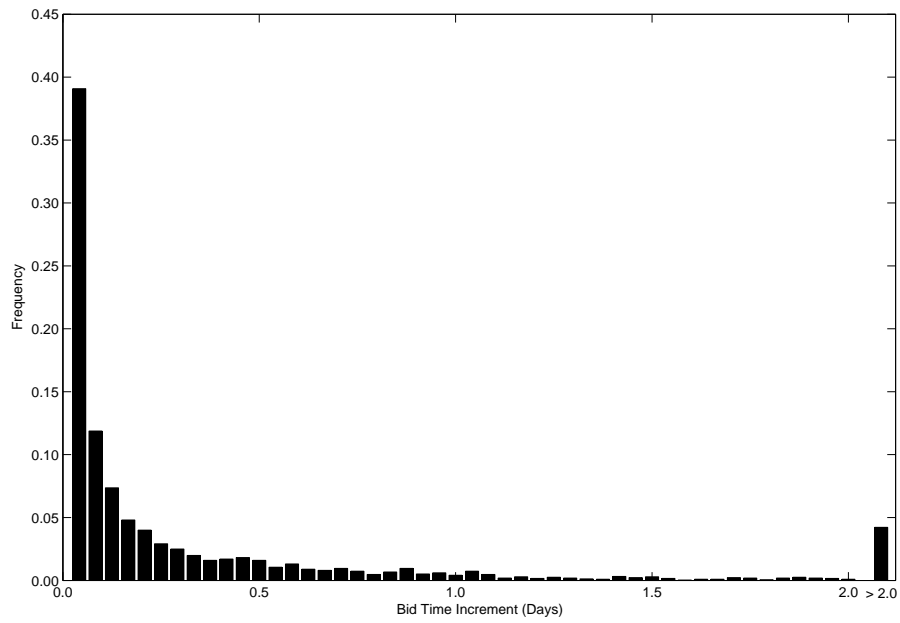


Figure 2: Histogram of Bid Time Increment

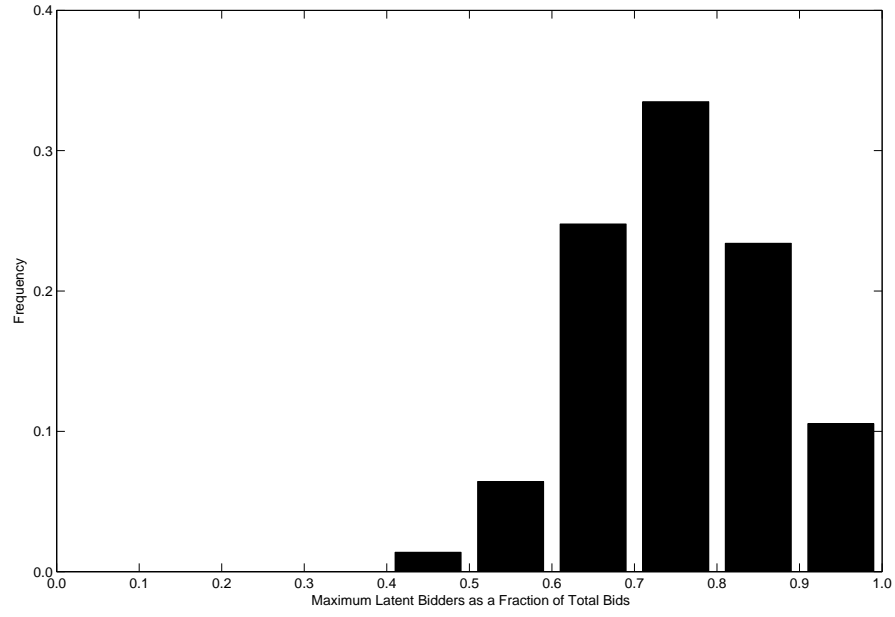


Figure 3: Histogram of Maximum Latent Bidders as a Fraction of Total Bids

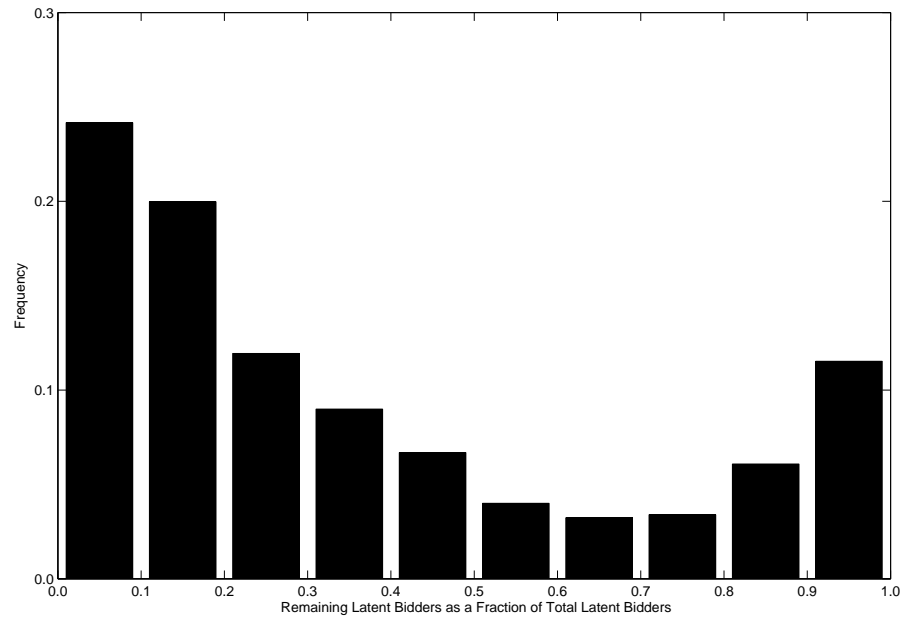


Figure 4: Histogram of  $\frac{\tilde{n}_{i,k}^*}{\tilde{n}_{i,k} + \tilde{n}_{i,k}^*}$

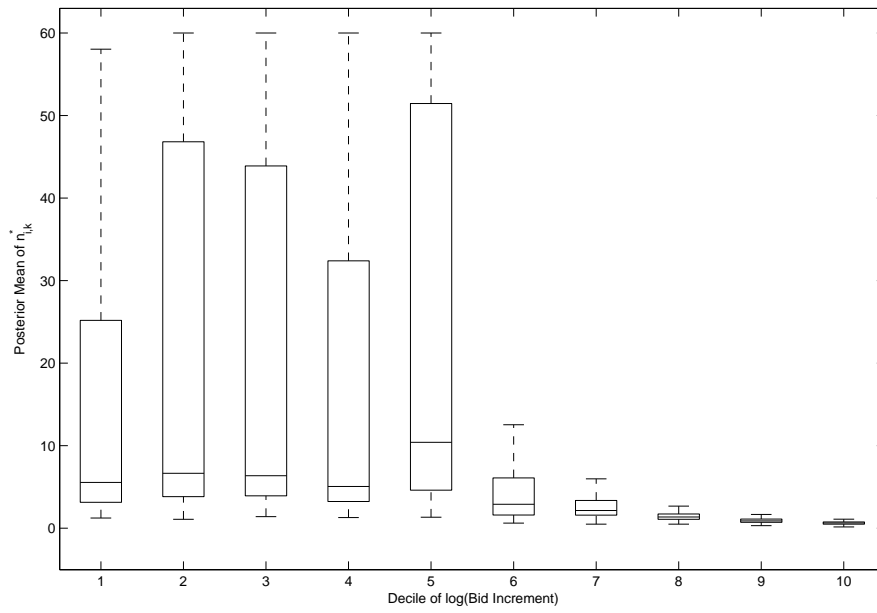


Figure 5: Posterior Means of  $n_{i,k}^*$  Based on Decile of  $\log(\text{Bid Increment})$

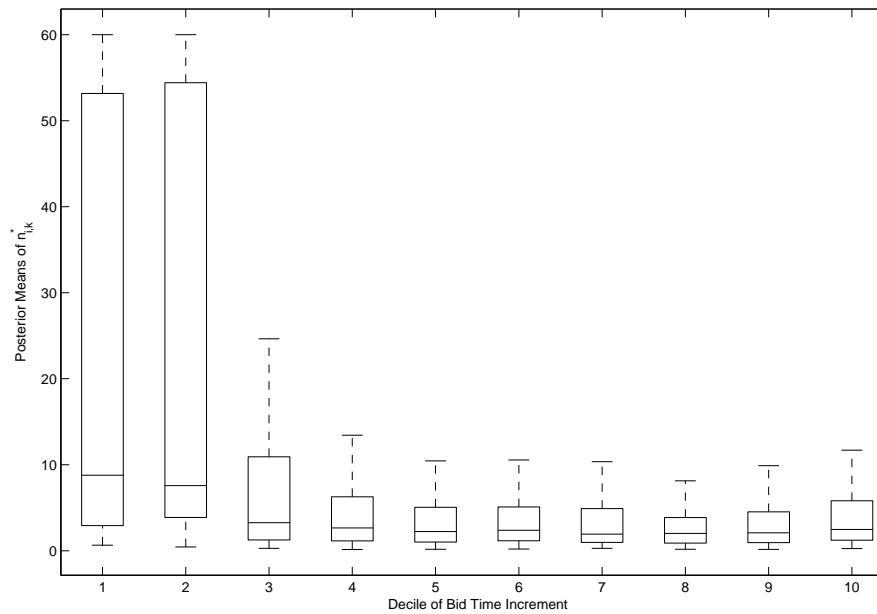


Figure 6: Posterior Means of  $n_{i,k}^*$  Based on Decile of Bid Time Increment

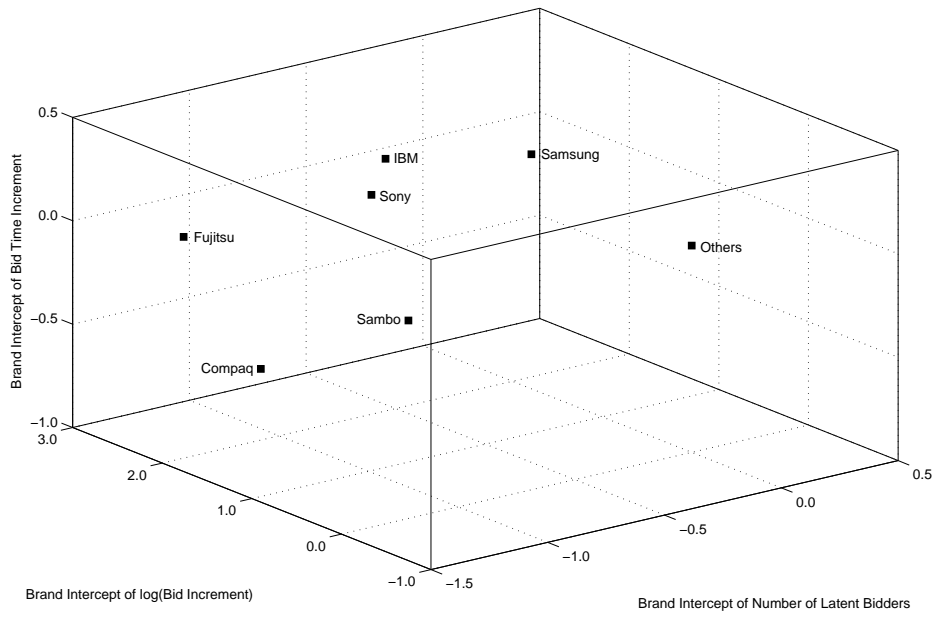


Figure 7: Brand Equity on Latent Bidders and Bidding Behavior