STATISTICAL METHODS IN ECOMMERCE RESEARCH

Chapter: Models of Bidder Activity Consistent with Self-Similar Bid Arrivals

Ralph P. Russo
Department of Statistics & Act. Science, The University of Iowa, Iowa City, IA 52242, USA

Galit Shmueli
Department of Decision and Information Technologies, R. H. Smith School of Business, University of Maryland, College Park, MD 20742, USA

Nariankadu D. Shyamalkumar
Department of Statistics & Act. Science, The University of Iowa, Iowa City, IA 52242, USA

A JOHN WILEY & SONS, INC., PUBLICATION

CHAPTER 1

MODELS OF BIDDER ACTIVITY CONSISTENT WITH SELF-SIMILAR BID ARRIVALS

Abstract: Bidder behavior is central to auction theory. In online auctions this behavior is largely hidden: while the bid placements of individual bidders is fully observable, bidder arrivals and departures, and bidder strategies are not. When aggregated, bid placements have been shown empirically to possess special properties. Let $N(s)$, $0 \le s \le T$, denote the bid arrival process associated with an online auction starting at time 0 and closing at time T. It has been observed in the literature that such a process becomes increasingly intense as the deadline approaches, and often displays a *self-similarity* property whereby the bid time distributions on the intervals $[s, T]$ become strikingly similar as s approaches T. In this chapter we identify a general process of bidder activity that (under appropriate conditions) generates a bid arrival sequence that possesses one or both of these properties.

1.1 INTRODUCTION

The fast growing popularity and importance of online auction websites such as eBay has led to a surge in empirical studies of a wide range of online auction phenomena. Using publicly available data from such sites, multiple studies have observed a divergence from the classic auction theory. Online auctions differ from *offline* auctions in several key respects: their length, their lowered barriers of entry for bidders and sellers, and their globalism. Phenomena such as bid *sniping* (last-minute bidding) and bid revising (an individual bidder placing multiple bids) are prevalent in online auctions where, according to the offline theory, they should not exist. Efforts to study online auctions have focused mainly on the identification and quantification of bidding strategies, and their justification from a game theoretic perspective [2, 3, 4, 5, 7, 9]. Our goal is to develop models of bidder activity consisting of bidder arrival and departure, and bid placement, that are consistent with the observable phenomena. The difficulty in creating such models stems from the fact that several aspects of bidder behavior (namely: bidder arrivals, departures, and strategies) are largely unobservable. Bid placements, on the other hand, are fully observable. In a recent paper, Shmueli et al. (2007) [8] proposed their BARISTA model which captures some of the main features of the bid arrival process of online auctions as observed and documented by several authors. Here we introduce a set of bidder behaviors that jointly produce bid arrivals that (under the appropriate conditions) give rise to BARISTA-like bid arrivals, or to bid arrival processes that possess some of the special characteristics that have been observed in empirical studies.

Let $N(s)$, $0 \leq s \leq T$, denote the bid arrival process associated with an online auction having a start time 0 and a deadline (hard close) T. In their study of online second-price auctions,[1] Roth & Ockenfels (2000) [6] noted two interesting characteristics common to the aggregations of such processes:

(P1) an *increasing intensity* of bid arrivals as the auction deadline approaches

(P2) a *striking similarity* in shape among the left truncated bid time distributions on $[s, T]$, as s approaches T.

The first phenomenon, also known as *bid sniping*, is a widely-known phenomenon in fixed-length auctions and has been documented in multiple empirical studies [e.g.,

[1] In a second-price auction the highest bidder wins the item and pays the second highest bid.

2, 3, 4, 8, 10]. The second phenomenon is referred to as *self-similarity*. We illustrate both in Figure 1.1., which displays the empirical cumulative distribution functions of normalized left truncated bid times based on 3643 bid times, placed in 189 online 7-day auctions on eBay.com for new Palm M515 PDAs. The graphs are plotted at several different resolution levels (or left truncation points), *zooming-in* from the last day, to the last 12 hours, to the last 3 hours, to the last 5 minutes. These functions have a similar shape, independent of scale, until the last few minutes (note the 5-minute curve) when the similarity breaks down (see [8]).

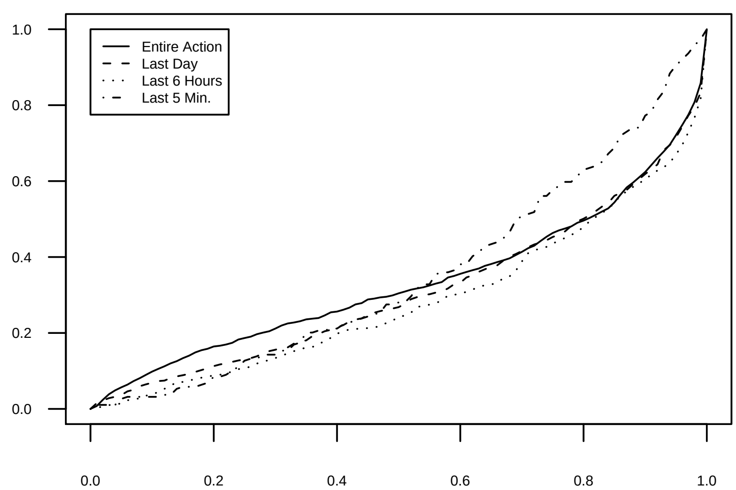

Figure 1.1. Empirical CDFs of Truncated Bid Arrivals - 189 Palm Auctions

For $0 \leq s < t \leq T$ define $N(s,t) = N(t) - N(s)$. *Increasing bid intensity* refers to the stochastic monotonicity of $N(t-\delta, t)$ as a function of t for any fixed $\delta > 0$, while *self-similarity* refers to the regularity in shape of the distribution functions

$$F_s(\eta) := \frac{N(T-\eta s, T)}{N(T-s,T)},\ 0 \leq \eta \leq 1$$

for all s sufficiently close to T (for $s \in [0, T-b]$, some b). Since $N(t-\delta, t)$ and $F_s(\eta)$ are determined empirically, to be precise we define these properties in terms of the expected bid counts: we say that the N-*process* has *increasing bid intensity* if $\mathbb{E}\left(N(t-\delta,t)\right)$ is increasing, and is *self-similar* over the interval $[b,T]$ if $E_s(\eta)$ defined as,

$$E_s(\eta) := \frac{\mathbb{E}\left(N(T-\eta s, T)\right)}{\mathbb{E}\left(N(T-s,T)\right)},\ \text{for } (s,\eta) \in [0, T-b] \times [0,1] \tag{1.1}$$

is independent of the value of s in $[0, T-b]$. In words, $E_s(\eta)$ is the same function of $\eta \in [0,1]$ for all $s \in [0, T-b]$. From the Cauchy equation (see p. 41 of [1]), for (1.1) to hold we must have for some $\gamma > 0$,

$$\frac{\mathbb{E}N(s)}{\mathbb{E}N(T)} = 1 - \left(1 - \frac{s}{T}\right)^{\gamma} \quad b \leq s \leq T, \tag{1.2}$$

in which case $E_s(\eta) = \eta^{\gamma}$ (see footnote 29 of [6]).

In the following section we define a *General Bid-Process* (GBP) of bidder arrivals and departures, and bid placement that (under sufficient conditions) yields a bid-arrival sequence that possesses one or both of the above properties. We derive an expression for the probability that an individual bidder is *active* at time s (has not departed the auction as of that time), and an expression for $\mathbb{E}\left(N(s)\right)$. Moreover, we show that $N(s)/N(T) \rightarrow \mathbb{E}(N(s))/\mathbb{E}(N(T))$ uniformly in $s \in [0, T]$ as the number of bidders increases. Under a simplifying restriction, the *General Bid-Process* reduces to the *General Poisson Bid-Process*, (GPBP), an aggregation of non-homogeneous Poisson processes having randomly determined start times and stopping rules. This latter process is related to the *BARISTA* process of Shmueli, et al (2007) [8], and is shown to possess the above property **(P1)** under very general conditions. A further simplification results in the *Self-Similar Bid-Process* (SSBP) which, as its name suggests, possesses the above property **(P2)**.

1.2 THE GENERAL BID-PROCESS

Auction theory focuses on bidder behavior and, in particular, on finding optimal bidding strategies for different auction formats. However, the online implementation of auctions has created a different environment where non-optimal bidding is often observed. Although various empirical studies have documented and quantified these phenomena, there exists a gap in the development of models of bidder behavior that are consistent with them. This owes largely to the fact that one can not directly observe bidder behavior in publicly available online auction data. Typically, bid placements are completely observable from the auction's bid history, whereas bidder arrivals and departures, and bidder strategies, are not. On eBay, for example, the temporal sequence of all bids placed over the course of the auction is publicly available. In particular, every time a bid is placed its exact time-stamp is posted. In contrast, the time when bidders first arrive at an auction is unobservable from the bid history. Bidders can browse an auction without placing a bid, and thereby not leave a trace or reveal their interest in that auction. That is, they can look at a particular auction, inform themselves about current bid- and competition-levels in that auction, and make decisions about their bidding strategies without leaving an observable trace in the bid history that the auction site makes public.

Our goal is to establish a model of bidding activity that is consistent with phenomena observable in the bid arrival process[2]. We now define bidder activity more formally.

Suppose that m bidders participate in an online auction starting at time 0 and closing at time T. The parameter m can be fixed or random (see Remark 3 below). With each bidder associate a random triple $\Theta = (X, \Pi, \mathcal{H})$, which we refer to as the bidder's *type*, that is comprised of an absolutely continuous random variable $X \in [0, T)$, a continuous function Π that maps $[0, T]$ into $(0, 1]$, and a family $\mathcal{H} = \{H_s\}$ of real valued distribution functions indexed by a real parameter $s \in [0, T]$, with $H_s(\cdot)$ having support $[s, T]$, where it is differentiable. The variable X represents the arrival time of an individual bidder into the auction, the function Π determines upon each of his bid placements whether he will remain in the auction and make a future bid, and the family $\mathcal{H}$ determines the timing of each bid that he makes. Given that the bidder's type is $\theta = (x, \pi, \{H_s\})$, he enters the auction at time x and places an initial bid at time $Y_1 \sim H_x$. If $Y_1 = y_1$, he departs the auction with probability $\pi(y_1)$, or otherwise places a second bid at time $Y_2 \sim H_{y_1}$. If $Y_2 = y_2$, he departs the auction with probability $\pi(y_2)$, or otherwise places a third bid at time $Y_3 \sim H_{y_2}$, etc., ultimately placing a random number of bids during $[0, T]$. All m bidders are assumed to act in a like manner, independently of each other.

In online auctions, *bid snipers* often attempt to place their final bid close to the deadline, so as to forestall a response from competing bidders. It has been observed that these late bids often fail to transmit due to technical reasons such as network

[2]eBay auctions use a proxy bidding mechanism, whereby bidders are advised to place the highest amount they are willing to pay for the auctioned item. The auction mechanism then automates the bidding process to ensure that the bidder with the highest proxy bid is in the lead at any given time. Thus, when Bidder A places a proxy bid that is lower than the highest proxy bid of (say) Bidder B, the new displayed highest bid and its time stamp will appear with Bidder B's username (although Bidder A is the one who placed the bid). In our discussion we shall consider such a bid as having been placed by A, rather than B, as it is A's action that led to a change in the displayed bid.

congestion. Thus, as $s \to T$, there may be a growing probability that a late placed bid fails to be recorded. This phenomenon can most efficiently be accommodated within our *GBP* by building the probability of a failed last bid directly into the π function as in (1.14).

For a more precise definition of the GBP, let $X_1, X_2, \ldots, X_m$ denote the arrival times of the m bidders. For $1 \le k \le m$, let $N_k(s)$ denote the number of bids placed by bidder k during the period $[0, s]$, $0 \le s \le T$, and let $Y_{k,j}$ denote the time of the j-th bid placed by the k-*th* bidder, $1 \le j \le N_k(T)$. For convenience, set $Y_{k,0} = X_k$, and suppose that:

(A1) $\Theta_1 = (X_1, \Pi_1, \mathcal{H}_1), \ldots, \Theta_m = (X_m, \Pi_m, \mathcal{H}_m)$ are independent, and identically distributed

(A2) $\Pr\left(Y_{k,j+1} \le t \mid Y_{k,j} = y_{k,j}\right) = H_{k,y_{k,j}}(t)$, for $1 \le k \le m$ and $j \ge 0$

(A3) The sequences $\{Y_{1,j}\}_{j \ge 0}, \ldots, \{Y_{m,j}\}_{j \ge 0}$ are independent

(A4) $\Pr\left[N_k(T) = r \mid Y_{k,r} = y_{k,r} \textit{ and } \Pi_k = \pi_k\right] = \pi_k(y_{k,r})$ for $r \ge 1$ and $1 \le k \le m$

Remark 1. We observe that condition **(A1)** allows dependence among the elements of Θ_k. Thus, the model can accommodate a tendency (say) for infrequent bidders to place their bid(s) late. We observe further that the above model accounts for heterogeneity in bidder probabilities of remaining in the auction after placing a bid, both across bidders and from bid to bid. Empirical evidence ([3]) suggests that a realistic distribution of Π should reflect the dichotomy of one-time bidders (*opportunists*) vs. multi-bid bidders (*participators*).

Remark 2. Since π is a continuous function on a compact set, it achieves its minimum ($= \pi_{\min}$) on $[0, T]$. Since π maps into $(0, 1]$, we have $\pi_{\min} > 0$. Thus, with each bid, the probability of departure is bounded below by $\pi_{\min}$. The total number of bids placed by the bidder on $[0, T]$ is therefore stochastically bounded above by a Geometric($\pi_{\min}$) variable, and thus has finite moments of all order.

Remark 3 (Poisson Bidder Arrivals). The parameter m can be either fixed or random. If random, it is natural to assume that m is Poisson distributed, as would be the case when bidders arrive in accordance with a non-homogeneous Poisson process having an intensity $\mu g(t)$, $t \in [0, T]$ where $\mu > 0$ and g is a density function on $[0, T]$. We then have that $X_1, \ldots, X_m$ is a random sample of random size $m \sim \text{Poisson}\,(\mu)$ from a fixed distribution with density function $g(t)$, $0 < t < T$.

1.2.1 A Single-Bidder Auction

The following two results pertain to an auction involving a single bidder ($m = 1$) whose *type* is $\theta = (x, \pi, \{H_s\})$. Let N_1 denote the resulting bid counting process. Obviously, a single bidder would not bid against himself. However, it is convenient to study the actions of a single bidder in the context of an m-bidder auction. We say that the bidder is *active* at time s if $N_1(T) > N_1(s)$ (he does not depart the auction during $[0, s]$). In particular, a bidder is active during the time period prior to his arrival. Let $p(s \mid \theta)$ denote the probability that this event occurs, and for $0 \leq s \leq t \leq T$ define

$$G_s(t \mid \theta) = \Pr\left(Y_{N_(s)+1} \leq t \mid \theta\right), \quad \text{on } \{N_1(T) > N_1(s);\ X < s\}$$

Given that the bidder arrived at time x (¡s), and is still active at time s, his next bid time $Y_{N_1(s)+1}$ has the distribution above. Let $g_s(\theta)$ denote the right derivative of $G_s(t \mid \theta)$ evaluated at s.

To derive the form of $g_s(\theta)$ we define the functional sequence

$$\begin{aligned}
\phi_0(s;t) &= h_s(t) \\
\phi_1(s;t) &= \int_s^t (1-\pi(u))\,\phi_0(s;u)\phi_0(u;t)du \\
&\vdots \\
\phi_n(s;t) &= \int_s^t (1-\pi(u))\,\phi_0(s;u)\phi_{n-1}(u;t)du
\end{aligned}$$

We now have,

$$g_s(\theta) = \frac{\sum_{n=0}^{\infty}\phi_n(x;s)}{\sum_{n=0}^{\infty}\int_s^T \phi_n(x;u)du}, \quad x < s$$

Proposition 1. *Suppose that*

$$\sup_{s \leq v \leq s+\delta} H_v(s+\delta) := \omega(s,\delta) \to 0 \textit{ as } \delta \to 0,\ \textit{all } s \in [0,T) \tag{1.3}$$

Then, for $0 \leq s \leq T$,

$$p(s \mid \theta) = 1_{x<s}\exp\left(-\int_x^s \pi(t)g_t(\theta)dt\right) + 1_{x>s}$$

Proof. If $x > s$ the result is trivial. Fix $s \in [x, T)$ and define

$$\pi_\delta = \inf\{\pi(t) : s \leq t \leq s+\delta\} \quad \text{and} \quad \pi^\delta = \sup\{\pi(t) : s \leq t \leq s+\delta\}$$

Writing $p(s)$ for $p(s \mid \theta)$ we have

$$p(s+\delta) \geq p(s)\left[1 - G_s(s+\delta \mid \theta) + G_s(s+\delta \mid \theta)(1-\pi^\delta)\left(1-\omega(s,\delta)\right)\right]$$

since an active bidder at time s remains active at time $s+\delta$ if either (i) $Y_{N_1(s)+1} \in (s+\delta, T)$, or (ii) $Y_{N_1(s)+1} < s+\delta$, the bidder stays active, and $Y_{N_1(s)+2} > s+\delta$. It follows that

$$\liminf_{\delta\to 0} \frac{p(s+\delta)-p(s)}{\delta} \geq -\pi(s)p(s)g_s(\theta)$$

Moreover,

$$p(s+\delta) \leq p(s)\left[1 - G_s(s+\delta \mid \theta) + G_s(s+\delta \mid \theta)(1-\pi_\delta)\right]$$

since an active bidder at time s is active at time $s+\delta$ only if (i) $Y_{N_1(s)+1} > s+\delta$, or (ii) $Y_{N_1(s)+1} < s+\delta$ and the bidder stays active. It follows that

$$\limsup_{\delta\to 0} \frac{p(s+\delta)-p(s)}{\delta} \leq -\pi(s)p(s)g_s(\theta)$$

Thus, $p'(s) = -\pi(s)p(s)g_s(\theta)$. Hence the proof. □

Remark 4. Condition (1.3) is a mild condition which will be assumed from here onwards. For the condition to fail would require a pathological $\mathcal{H}$.

The following is an easy consequence of the above and is hence stated without proof:

Proposition 2. *For* $0 \leq s \leq T$*,*

$$\mathbb{E}\left[N_1(s) \mid \theta\right] = 1_{x<s} \int_x^s p(t \mid \theta) g_t(\theta) \mathrm{d}t$$

1.2.2 A Multi-Bidder Auction

Returning to the case of m bidders, we state a uniform limit result for $N(s)/N(T)$. It should be noted that many online auctions involve small numbers of bidders. *Self-similarity*, as studied in [6], is not a phenomenon that can be easily observed from an individual auction where the number of bids are few. To observe the *self-similarity* property one usually must aggregate many *equivalent* auctions (that is, auctions for the same item, over the same duration, etc.). Such aggregations often involve hundreds of bidders, and hence our interest in $m \to \infty$.

Proposition 3. *If* $\mathbb{E}\left(N_1(T)\right) < \infty$*, then as* $m \to \infty$*,*

$$\sup_{0\leq s\leq T} \left| \frac{N(s)}{N(T)} - \frac{\mathbb{E}\left(N_1(s)\right)}{\mathbb{E}\left(N_1(T)\right)} \right| \to 0 \ \textit{ almost surely}$$

Proof. For fixed $s \in [0,T]$, $N(s)$ is the sum of m independent and identically distributed random variables. By the *strong law of large numbers* we have almost sure pointwise (in s) convergence:

$$\frac{N(s)}{N(T)} \to \frac{\mathbb{E}\,(N_1(s))}{\mathbb{E}\,(N_1(T))} \ \textit{almost surely}$$

By the continuity of the limit as a function of s, and by Polya's Theorem, the above convergence is uniform in s. Hence the proof. □

Remark 5. By Remark 2 we have $\mathbb{E}\,(N_1(T)|\theta) \leq 1/\pi_{\min}$, so that by the double expectation formula, $\mathbb{E}\,(N_1(T)) \leq \mathbb{E}\,(1/\Pi_{\min})$ where $\Pi_{\min} = \min\{\Pi(s) : s \in [0,T]\}$. Thus, the finiteness of $\mathbb{E}\,(1/\Pi_{\min})$ is sufficient for the convergence in Proposition 3.

Remark 6 (Poisson bidder arrivals). In the case where bidders arrive in accordance with a non-homogeneous Poisson process with intensity $\mu g(t)$ (as in Remark 3), Proposition 3 and a standard coupling argument yields the following result:

$$\varepsilon > 0 \text{ and } \mathbb{E}N_1(T) < \infty \Rightarrow \lim_{\mu\to\infty} \Pr\left(\sup_{0<s<T}\left|\frac{N(s)}{N(T)} - \frac{\mathbb{E}N_1(s)}{\mathbb{E}N_1(T)}\right| > \varepsilon\right) = 0 \qquad (1.4)$$

The practical significance of (1.4) is that in an auction with a high number of participants, or in aggregations of many equivalent auctions, the observed distribution of bid times on $[0,T]$ will be uniformly close to the deterministic function $\mathbb{E}N_1(s)/\mathbb{E}N_1(T)$ with high probability.

1.3 THE GENERAL POISSON BID-PROCESS

We consider now a simplifying restriction on the form of $\mathcal{H}$. Given $\theta = (x, \pi, \{H_s\})$, suppose that all of the H_s functions are driven by the single (randomly determined) function H_0 as follows:

$$H_s(t) = \frac{H_0(t) - H_0(s)}{1 - H_0(s)} \text{ for } 0 \leq s \leq t < T \qquad (1.5)$$

Under condition (1.5) the bidder's *j-th* bid time ($j \geq 1$), conditional on $Y_{j-1} = y_{j-1}$, is distributed as H_0 restricted to $[y_{j-1}, T]$. We note that a bid time chosen from H_s, that is not realized by time $t > s$, is probabilistically equivalent to one chosen from H_t.

In this subsection, N_1 denotes the single-bidder process of bid arrivals under condition (1.5). We define an auxiliary bid counting process $M(s)$, $0 \leq s \leq T$, under the additional conditions that $\Pr(X = 0) = 1$ (the bidder enters the auction at time $s = 0$) and $\Pr(\Pi(s) \equiv 0) = 1$ (the bidder never departs the auction). We observe that: $M(0) = 0$ and that the M-process possesses independent increments. Moreover, writing H for H_0 we have

$$\begin{aligned}
&\limsup_{\delta\to 0} \frac{\Pr(M(s+\delta) - M(s) \geq 2)}{\delta} \\
&\qquad\qquad \leq \limsup_{\delta\to 0} \left[\frac{H(s+\delta) - H(s)}{1 - H(s)}\right]^2 \frac{1}{\delta} \\
&\qquad\qquad = \left(\frac{1}{1-H(s)}\right)^2 h(s) \limsup_{\delta\to 0} \left[H(s+\delta) - H(s)\right] = 0 \qquad (1.6)
\end{aligned}$$

and by (1.6),

$$\begin{aligned}
\lim_{\delta\to 0} \frac{\Pr(M(s+\delta) - M(s) = 1)}{\delta} &= \lim_{\delta\to 0} \left[\frac{\Pr(M(s+\delta) - M(s) \geq 1)}{\delta}\right] \\
&= \lim_{\delta\to 0} \frac{H(s+\delta) - H(s)}{(1 - H(s))\,\delta} \\
&= \frac{h(s)}{1 - H(s)}
\end{aligned}$$

The M-process is thus a non-homogeneous Poisson process with intensity function

$$\lambda(s) = \frac{h(s)}{1 - H(s)}$$

Intuitively, when the auction clock reaches time s, no matter how many bids have been placed, and no matter when they have been placed, the probability that a bid will be placed during $(s, s + \delta)$ is approximately $\delta h(s)/(1 - H(s)$. Given θ (associated with the N_1-process) we may use the function $\pi(\cdot)$ to randomly label an *M-process* arrival at time s as A or B with respective probabilities $1 - \pi(s)$ and $\pi(s)$. The resulting offspring processes M_A and M_B are independent non-homogeneous Poisson processes with respective intensity functions

$$\lambda_A(s) = \frac{(1 - \pi(s))h(s)}{1 - H(s)} \quad \text{and} \quad \lambda_B(s) = \frac{\pi(s)h(s)}{1 - H(s)}$$

Note that arrivals from the *N_1-process* are the M arrivals that occur after the bidder arrival time X, up to and including the first arrival from M_B. That is, N_1 is a nonhomogeneous Poisson process with intensity function λ, restricted to the random interval $[X,\ T]$, and stopped upon the first arrival from M_B. The N-process (involving all m bidders) is an aggregation of m such independent processes.

Given $\theta = (x, \pi, \{H_s\})$, a bidder is active at time s if and only if either $x > s$, or $x < s$ and there are no arrivals from M_B during the period $[x, s]$. Accordingly,

$$\begin{aligned} p(s \mid \theta) &= 1_{x<s} \Pr\left[M_B(s) - M_B(x) = 0\right] + 1_{x \geq s} \\ &= 1_{x<s} \Pr\left[Poisson\left(\int_x^s \lambda_B(t)dt\right) = 0\right] + 1_{x \geq s} \qquad (1.7) \\ &= 1_{x<s} \exp\left[-\int_x^s \frac{\pi(t)h(t)}{1-H(t)}dt\right] + 1_{x \geq s} \end{aligned}$$

From the above we obtain the conditional bid intensity $\lambda(\cdot|\theta)$ of an individual bidder of type θ:

$$\lambda(s|\theta) = \begin{cases} \exp\left[-\int_x^s \frac{\pi(t)h(t)}{1-H(t)}dt\right] \frac{h(s)}{1-H(s)}, & s > x; \\ \\ 0, & s \leq x. \end{cases}$$

In the case where $\limsup_{s \to T} \pi(s) < 1$ and $\lim_{s \to T} h(s) > 0$, the conditional intensity explodes as s approaches T, *i.e.* $\lim_{s \to T} \lambda(s|\theta) = \infty$. The condition on $\pi(\cdot)$ can be dropped if $h(s)$ increases to infinity sufficiently fast as s approaches T (e.g., at any polynomial rate).

1.3.1 A Constant Probability of Departure

In this subsection we suppose that (upon each bid placement) the bidder has a randomly determined time invariant probability of departure:

$$\Pr\left(\Pi(s) = \Pi(0),\ 0 \leq s \leq T\right) = 1 \qquad (1.8)$$

Under the above assumption, the number of bids placed by an individual bidder is geometrically distributed with a randomly determined parameter. A time invariant departure probability, while certainly not true for the entire auction duration, should be approximately true over short intervals. Under (1.5) and (1.8), statement (1.7) reduces to

$$p(s \mid \theta) = 1_{x<s} \left[\frac{1-H(s)}{1-H(x)}\right]^{\pi(0)} + 1_{x>s}. \qquad (1.9)$$

By Proposition 2 we get

$$\begin{aligned}
\mathbb{E}\left(N_1(s) \mid \theta\right) &= 1_{x<s} \int_x^s \left[\frac{1-H(t)}{1-H(x)}\right]^{\pi(0)} \frac{h(t)}{1-H(t)} dt \\
&= \frac{1_{x<s}}{\pi(0)} \left(1 - \left[\frac{1-H(s)}{1-H(x)}\right]^{\pi(0)}\right)
\end{aligned}$$

so that by the double expectation formula, writing Π for $\Pi(0)$,

$$\mathbb{E}\left(N_1(s)\right) = \mathbb{E}\left(\frac{1_{X<s}}{\Pi}\left(1 - \frac{1-H(s)}{1-H(X)}\right)^{\Pi}\right) \tag{1.10}$$

Hence, by Proposition 3, under conditions (1.8) and (1.5), if $\mathbb{E}\Pi^{-1} < \infty$,

$$\sup_{0<s<T} \left| \frac{N(s)}{N(T)} - \mathbb{E}\frac{1_{X<s}}{\Pi}\left(1 - \frac{1-H(s)}{1-H(X)}\right)^{\Pi} \frac{1}{\mathbb{E}\Pi^{-1}} \right| \to 0 \quad a.s. \text{ as } m \to \infty$$

and for Poisson arrivals (as in Remarks 3 and 6),

$$\lim_{\mu\to\infty} \Pr\left(\sup_{0<s<T} \left| \frac{N(s)}{N(T)} - \mathbb{E}\frac{1_{X<s}}{\Pi}\left(1 - \frac{1-H(s)}{1-H(X)}\right)^{\Pi} \frac{1}{\mathbb{E}\Pi^{-1}} \right| > \varepsilon\right) = 0$$

1.3.2 The Self-similar bid-process

Suppose that conditions (1.5) and (1.8) hold, and that for some constant $r > 0$ (the same for all bidders), we have

$$H_0(s) = 1 - \left(1 - \frac{s}{T}\right)^{r/\pi(0)} \tag{1.11}$$

Suppose, in addition, that all bidders arrive by time $b < T$:

$$\Pr\left(0 \le X < b\right) = 1 \tag{1.12}$$

Under (1.11), the higher a bidders likelihood of departure upon the placement of a bid at time s, the stochastically greater the time of his next bid (i.e., the more inclined he is to choose his next bid near the deadline T). Condition (1.11) ties the selection function H_0 directly to the constant departure probability $\pi(0)$. Again writing Π for $\Pi(0)$ and assuming that $\mathbb{E}\Pi^{-1} < \infty$, we have by (1.10)

$$\begin{aligned}
\mathbb{E}N_1(T-s,T) &= \mathbb{E}\left(\Pi^{-1}\left(\frac{s}{T-X}\right)^r 1_{X<T-s} + \Pi^{-1}1_{X>T-s}\right) \\
&= \mathbb{E}\Pi^{-1}\left(\frac{1}{T-X}\right)^r s^r \quad for\ s < T-b
\end{aligned}$$

Thus, for $(s, \eta) \in [0, T-b] \times [0, 1]$, we have

$$E_s(\eta) = \eta^r, \tag{1.13}$$

where $E_s(\eta)$ is defined in (1.1)

By statement (1.13) and Proposition 3 we have for large m

$$N_s(\eta) = \frac{N(T-\eta s, T)}{N(T-s, T)} \approx \eta^r, \ \text{ for } (\eta, s) \in [0,1] \times [0, T-b].$$

1.3.3 The BARISTA Process

The efforts of bidders to place bids late in the auction are often thwarted by transmission failures. We build this phenomenon into the *GPBP* by assuming (1.11) and supposing that each bidder has a randomly determined time invariant probability of departure upon the placement of all bids on $[0, T-d]$, for some small $d > 0$, and that this probability is magnified by a constant β for all bids placed after time $T-d$ (β and d being the same for all bidders):

$$\Pi(s) = \Pi(0) 1_{s \leq T-d} + \beta \Pi(0) 1_{s > T-d} \tag{1.14}$$

The resulting process of bid arrivals is a *GPBP* (but not a *SSBP*) as described in section 1.3.1, and can thus also be characterized as an aggregation of independent non-homogeneous Poisson processes with randomly determined start times and stopping rules. At time $s \in (b, T-d]$, the bid intensity associated with an individual bidder of type θ is

$$\begin{aligned}
\lambda_1(s \mid \theta) &= p(s \mid \theta) \frac{h(s)}{1-H(s)} \\
&= \left(\frac{1-H(s)}{1-H(x)}\right)^{\pi(0)} \frac{h(s)}{1-H(s)} \quad \text{by (1.9)} \\
&= \left(1 - \frac{s}{T}\right)^{r-1} \left(1 - \frac{x}{T}\right)^{-r} \frac{r}{\pi(0) T} \quad \text{by (1.11)}
\end{aligned}$$

and hence, for a given collection of m bidder types, the intensity of the bid arrival sequence at $s \in (b,\, T-d]$ is given by

$$\lambda(s \mid m,\, \theta_1, \ldots,\, \theta_m) = \left[\frac{r}{T} \sum_{k=1}^{m} \frac{1}{\pi_k(0)} \left(1 - \frac{x_k}{T}\right)^{-r}\right] \left(1 - \frac{s}{T}\right)^{r-1}$$

For $s \in (T-d, T]$,

$$\begin{aligned}\lambda_1(s \mid \theta) &= p(T-d \mid \theta)\left(\frac{1-H(s)}{1-H(T-d)}\right)^{\beta\pi(0)} \frac{h(s)}{1-H(s)} \\ &= \left(\frac{d}{T}\right)^{r-r\beta}\left(1-\frac{x}{T}\right)^{-r} \frac{r}{\pi(0)T}\left(1-\frac{s}{T}\right)^{r\beta-1}\end{aligned}$$

and hence, for a given collection of m bidder types, the intensity of the bid arrival sequence at $s \in (b, T)$ is given by

$$\lambda(s \mid m,\ \theta_1, \ldots,\ \theta_m) = \begin{cases} c\left(1-\frac{s}{T}\right)^{r-1}, & s \in (b, T-d] \\ \\ c\left(\frac{d}{T}\right)^{r-r\beta}\left(1-\frac{s}{T}\right)^{r\beta-1}, & s \in (T-d, T] \end{cases}$$

where

$$c = \left[\frac{r}{T}\sum_{k=1}^{m} \frac{1}{\pi_k(0)}\left(1-\frac{x_k}{T}\right)^{-r}\right]$$

is the result of the realization $(m, \theta_1, \ldots, \theta_m)$. This is the form of the two-stage BARISTA [8] intensity with $d_1 = 0$, $d_2 = d$, $\alpha_2 = r$ and $\alpha_3 = \beta r$. It can be shown similarly that multiple shifts in the departure probability will lead to an intensity of the multi-stage BARISTA form. In particular a double shift yields the 3-stage intensity discussed in [8].

Remark 7. We note that in the case of $r\beta < 1$, $\lambda(s \mid m,\ \theta_1, \ldots,\ \theta)$ as a function of s is increasing to infinity (*i.e.* increasing and exploding) as s approaches T.

Working with the general Poisson model under condition (1.8), we now derive the distribution of an individual bidder's final bid time Y_{final}. Again, writing Π for $\Pi(0)$, we have by (1.9),

$$P\left(Y_{final} > s\right) = \mathbb{E}\left(\frac{1-H(s)}{1-H(X)}\right)^{\Pi} 1_{X<s} + \Pr\left(X > s\right) \tag{1.15}$$

and thus

$$\mathbb{E}Y_{\text{final}} = \mathbb{E}X + \int_0^T \mathbb{E}\left(\frac{1-H(s)}{1-H(X)}\right)^{\Pi} 1_{X<s}\mathrm{d}s$$

For the *Self-similar bid process* the above statements simplify to:

$$P\left(Y_{final} > s\right) = \mathbb{E}\left(\frac{T-s}{T-X}\right)^{r} 1_{X<s} + \Pr\left(X > s\right)$$

and

$$\mathbb{E}Y_{\text{final}} = \mathbb{E}X + \int_0^T \mathbb{E}\left[\left(\frac{T-s}{T-X}\right)^{r} 1_{X<s}\right]\mathrm{d}s$$

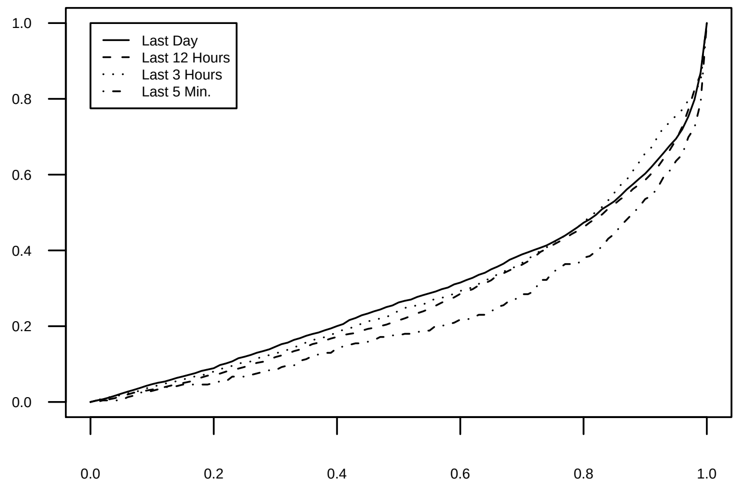

Figure 1.2. Empirical CDFs of Truncated Bid Arrivals - Simulated Data

In the case of m-bidders, $Y_{m,\text{final}}$ is the sample maximum from a random sample of size m from the distribution of Y_{final}. Hence, properties of $Y_{m,\text{final}}$ are easily derived from those of Y_{final} given above. Also, by using conditioning we can extend the above results to the case of random m.

1.3.4 Examples

We end this section with two simulations.

Example 1 (SSBP). We set $m = 1000$, $T = 7$, $X \sim U(0, 5.6)$, $\Pr(\Pi(0) \leq t) = 4t^2$ for $t \in (0, 1/2)$ and $r = 2/5$. In our simulation, 1000 bidders arrive uniformly during the first 5.6 days (the first 80%) of a 7 day auction, and their departure probabilities constitute a random sample from a distribution having a triangular shape on $(0, 1/2)$, so that each bidder is expected to place 4 bids (4619 were placed in our simulation). Moreover, any bidder placing a bid at time t and opting to remain active, has

$$\Pr(\text{Time of Placement of Next Bid} > s) = \left[\frac{1-s}{1-t}\right]^{\frac{5}{2\pi}}.$$

Uniform bidder arrivals during a fixed time span is a natural assumption. The exact distribution of the arrival times over $(0, 5.6)$ determines the distribution of the number of active bidders at time point 5.6, but is otherwise irrelevant. In the

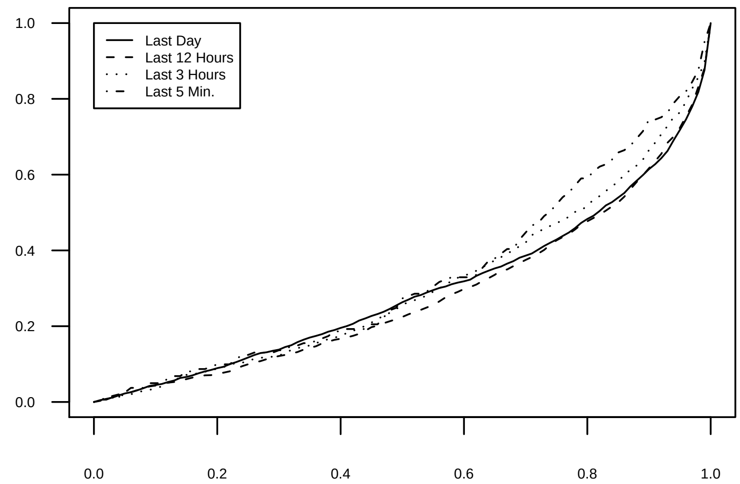

Figure 1.3. Empirical CDFs of Truncated Bid Arrivals - Simulated Data

case of U$(0, 5.6)$, we expect $.678m = 678$ bidders to be active at time $t = 5.6$. In our simulation, the observed number of such bidders was slightly lower (665). The function H is of the form required in the SSBP model. Our value of $r = 2/5$ is in the range $(0, 1)$ which guarantees that h increases as t approaches T. Figure 1.2. displays the empirical cumulative distribution functions of normalized left truncated bid times. Note the similarity (excluding the five-minute graph) of Figures 1.1. and 1.2. .

While the simple model of the above example is able to capture the self similarity displayed in the Palm data, it fails to capture its breakdown as indicated by the 5-minute curve of Fig. 1.1. In the following example we capture this breakdown by magnifying, during the final minute, each bidder's probability of departure (as in (1.14)). This magnification incorporates the greater likelihood of nontransmittal of bids during the final minute.

Example 2 (GPBP, BARISTA). We maintain the same set-up used in the above example except that $\pi(0)$ is doubled during the final minute ($\beta = 2$ in (1.14)). Figure 1.3. displays the empirical cumulative distribution functions of normalized left truncated bid times. Note the similarity of Figures 1.1. and 1.3..

REFERENCES

1. J. Aczel. *Lectures on Functional Equations and Their Applications*. Academic Press, New York, 1966.
2. P Bajari and A Hortacsu. *New Economy Handbook*, chapter Cyberspace Auctions and Pricing Issues: A Review of Empirical Findings. Elsevier, 2003.
3. Ravi Bapna, Paulo Goes, Alok Gupta, and Yiwei Jin. User heterogeneity and its impact on electronic auction market design: An empirical exploration. *MIS Quarterly*, 28:1:21–43, 2004.
4. S. Borle, P. Boatwright, and J. B. Kadane. The timing of bid placement and extent of multiple bidding: An empirical investigation using ebay online auctions. *Statistical Science*, 21 (2):194–205, 2006.
5. A Ockenfels and A E Roth. The timing of bidding in internet auctions: Market design, bidder behavior and artificial agents. *AI Magazine*, 23(3):79–87, 2002.
6. A E Roth and A Ockenfels. Last-minute bidding and the rules for ending second-price auctions: Theory & evidence from a natural experiment on the internet. Technical report, NBER Working Paper #7729, 2000.
7. A E Roth and A Ockenfels. Last-minute bidding and the rules for ending second-price auctions: Evidence from ebay and amazon auctions on the internet. *The American Economic Review*, 92(4):1093–1103, 2002.
8. G. Shmueli, R. P. Russo, and W. Jank. The barista: A model for bid arrivals in online auctions. *Annals of Applied Statistics*, 2007 (in press).
9. H. Varian. Online auctions as a laboratory for economists to test their theories. *New York Times*, Nov. 16, 2000.
10. R T Wilcox. Experts and amateurs: The role of experience in internet auctions. *Marketing Letters*, 11(4):363–374, 2000.