Dynamic Profiling of Online Auctions Using Curve Clustering

Wolfgang Jank & Galit Shmueli

Department of Decision and Information Technologies
The Robert H. Smith School of Business
University of Maryland

Abstract

Electronic commerce, and in particular online auctions, have received an extreme surge of popularity in recent years. While auction theory has been studied for a long time from a game-theory perspective, the electronic implementation of the auction mechanism poses new and challenging research questions. Although the body of empirical research on online auctions is growing, there is a lack of treatment of these data from a modern statistical point of view. In this work, we present a new source of rich auction data and introduce an innovative way of modelling and analyzing online bidding behavior. In particular, we use functional data analysis to investigate and scrutinize online auction dynamics. We describe the structure of such data and suggest suitable methods, including data smoothing and curve clustering, that allow one to profile online auctions and display different bidding behavior. We illustrate the methods on a set of eBay auction data and tie our results to the existing literature on online auctions.

Key words and phrases: functional data analysis, smoothing, penalized splines, clustering, unsupervised classification, k-medoids, electronic commerce, online auction, eBay, bidding behavior, bid sniping, bid shilling

1 Introduction

The public nature of many online marketplaces has allowed empirical researchers new opportunities to gather and analyze data. One example of an online marketplace is the online auction. Online auctions have become a very popular way for both businesses and consumers to exchange goods. One of the biggest online marketplaces and currently the biggest C2C online auction place is *eBay* (`http://www.ebay.com`). In 2002, eBay had more than 69 million registered users and it offered more than 16 million items on its site. Since eBay archives detailed records of its completed auctions, it is a great source for immense amounts of high quality data.

There has been an extensive amount of research on classical auction theory (see Milgrom & Weber, 1982; Klemperer, 1999, for an introduction and overview). Classical auction theory typically centers around the analysis of optimal auctions and the effects of relaxing some of their assumptions. While there has been some empirical research in the area (e.g. Hendricks & Paarsch, 1995), the scarce involvement of statisticians in the field is most likely due to the absence of widely available data. However, more and more bidding-data are now becoming available thanks to the recent surge of online auctions and the capability of collecting data conveniently over the internet. While online auctions have the advantage of increasing the understanding of classical auction theory, they also pose new research questions. Indeed, recent research suggests the need for additional investigations of the impact of the electronic implementation of the auction mechanism (e.g. Klein & O'Keefe, 1999). For instance, empirical studies provide evidence that classical auction theory may not generally carry over to the online context (Lucking-Reiley, 1999). Empirical work has also observed the prevalence of "bid sniping," in which the majority of bids arrive in the final moments of the auction (Roth & Ockenfels, 2002). Fraudulent behavior has also become more common due to the anonymity of the internet. "Bid shilling," for example, is a fraudulent practice in which the seller bids on his/her own auction in order to increase the price (e.g. Kauffman & Wood, 2003). In this work, we address some of these research questions with the help of a set of novel statistical tools, often referred to as *functional data analysis*.

The focus and goal of our paper is to frame online auction data in a way suitable for statistical reasoning. Auction data typically arrive in the form of a sequence of bids placed over a period of time between the start and end of the auction. This is often referred to as the "bid history" of the auction. In this work, we view the bid history as characteristic of the bidding dynamics of the auction. We assume that the observed bids are realizations of an underlying (but unobserved) continuous bidding curve. Each bidding curve is unique to its auction. Variation between two bidding curves is due to a larger population of heterogeneous bidding dynamics. Thus, the goal is to use functional data analysis to detect regularities and patterns within a sample of bidding curves drawn from this population. Furthermore, we show not only how the statistical methods can be applied to the data, but also how the different steps translate into the domain of online auctions.

The statistics literature is currently experiencing a tremendous amount of interest in methodology as well as applications related to the analysis of functional data. While many of the underlying ideas have been around for a longer time, the name functional data analysis is typically attributed to the work of Ramsay & Silverman (1997). In functional data analysis, the object of interest is often a set of curves, shapes, objects, or, more generally, a set of *functional observations*, rather than a set of data points, as it is typically the case in classical statistics. There are a number of recent articles devoted to the generalization of standard statistical methodology to the context of functional observations. Fraiman & Muniz (2001), for instance, develop a measure of centrality for a given functional observation within a group of curves. James et al. (2000) present a technique for principal component analysis of a set of sparsely-sampled curves (see also Ocana et al., 1999). Other exploratory tools have been developed, among them tools for curve-clustering (see Abraham et al., 2003; James & Sugar, 2003; Tarpey & Kinateder, 2003) as well as curve-classification (see Hall et al., 2001; James & Hastie, 2001). Classical statistical methods have also been generalized to functional canonical correlation analysis (He et al., 2003) as well as functional ANOVA (Fan & Lin, 1998; Guo, 2002). Modelling functional observations is another field of particular interest. While Faraway (1997) discusses regression analysis when the response arrives in functional form (see also Yu & Lambert, 1999; Cuevas et al., 2002; Ratcliffe et al., 2002b), James (2002) considers generalized linear models for functional data (see also Ratcliffe et al., 2002a). Ramsay (2000), on the other hand, fits differential equation models to data of functional form (see also Ramsay & Ramsey, 2002). While this list is far from complete, it shows some of the current methodological efforts in this emerging field.

Functional data analysis has been applied to many areas in which statisticians are actively involved, such as the agricultural sciences (Ogden et al., 2002), the behavioral sciences (Rossi et al., 2002) as well as medical research (Pfeiffer et al., 2002). The method has been used to analyze the dynamics of seasonally-varying production indices (Ramsay & Ramsey, 2002) as well as to predict El Niño (Besse et al., 2000). However, while there exist many more applications in which functional data methods have been fruitful, it appears that this set of tools has not yet been applied to the analysis of bidding behavior or to data originating from electronic commerce.

While bidding behavior has been studied extensively, especially in the field of economics, it has not received a widespread popularity in the statistics literature. Presumably, one reason for this is the previous lack of publicly available bidding data. One possible source of auction data is the government. For instance, Pelto (1971) considers bidding data for offshore oil and mineral rights obtained from the U.S Department of the Interior. Other government data sources are the Federal Communications Commission or the Treasury Department (see `http://www.firstgov.gov/shopping/ auctions/auctions.shtml` for a listing of all government auctions). One drawback of these data sources is the rather limited amount of information that is revealed. Li et al. (2003), who also study offshore mineral rights, report that their data only contain the number of bidders and the bid amount. While other data sources exist, they are often harder to find. Elyakime et al. (1997), for instance, study auctions for standing timber in France but do not report the data.

The increasing popularity of the internet and electronic commerce has made online auction houses like *eBay*, *Amazon* or *uBid* a popular place for exchanging goods. Moreover, the wealth of available information makes these places an ideal ground for investigating bidding behavior. Indeed, eBay reports for each closed auction not only the number of bids placed and their values but also information about the temporal sequence of the bids, the auction participants, the product features, and the auction options. eBay's popularity and its data-accessibility have made it a center of recent research efforts (see Bajari & Hortacsu, 2003; Klein & O'Keefe, 1999; Lucking-Reiley, 1999, 2000; Roth & Ockenfels, 2002). However, while it is possible to collect data from online auction houses "manually" by directing one's internet browser to the corresponding web page, collecting a large amount of data in this way can be burdensome.

Modern technologies allow for a convenient collection of large amounts of high quality data from the internet. The use of *web agents* or *web spiders* facilitate the creation of large databases of bidding data. A web agent is a software application, typically based on a programming language like *Pearl* or *Java*, that "crawls" over an internet site or a collection of web pages and gathers the desired information. In this form, data on several hundred auctions can be collected in a matter of only minutes, ready to be explored by the statistician!

This paper is organized as follows. In Section 2 we describe the data and the mechanism that generates and collects them. In Section 3 we discuss the functional representation of an online auction and the assessment of bidding dynamics. Section 4 proposes a method for profiling online auctions based on unsupervised curve-clustering. In Section 5 we apply the methods to a collection of Palm Pilot auctions and discuss the results. The paper concludes with final remarks in Section 6.

2 Auction structure and available data on eBay.com

2.1 How eBay's auctions work

The auction format typically found on eBay is called "proxy bidding". It is a variant of the second price sealed-bid auction ("Vickrey auctions") (see e.g. Krishna, 2002), where individuals submit a "proxy bid", which is the maximum value they are willing to pay for the item. The auction mechanism automates the bidding process to ensure that the person with the highest proxy bid is in the lead of the auction. The winner is the highest bidder and pays the second highest bid. For example, suppose that bidder A is the first bidder to submit a proxy bid on an item with a minimum bid of $10 and a minimum bid-increment of $0.50. Suppose that bidder A places a proxy bid of $25. Then eBay's web page automatically displays A as the highest bidder, with a bid of $10. Next, suppose that bidder B enters the auction with a proxy bid of $13. eBay still displays A as the highest bidder, however it raises the displayed high-bid to $13.50, one bid increment above the second-highest bid. If another bidder submits a proxy bid above $25.50, bidder A is no longer in the lead. However, if bidder A wishes, he or she can submit a new proxy bid. This process continues until the auction ends. Unlike other auctions, eBay has strict ending times, ranging between 3 and 10 days from the opening of the auction, as determined by the seller. While eBay discloses only the time of the bid and the username of the bidder during the live auction, it posts the complete bid histories of closed auctions for a duration of at least 30 days on its web site (see `http://listings.ebay.com/pool1/ listings/list/completed.html`).

2.2 eBay's bid history data

Figure 1 shows a typical example of bid histories found on eBay's web site. The top of Figure 1 displays a summary of the information relevant to the auction. In particular, after providing a description of the item and the item number, it displays the current bid (which is identical to the winning bid since the auction is closed), the starting bid (which is the minimum bid), the quantity of items auctioned (which is always equal to one in the "proxy bid" auction type), and the number of bids received during the auction. It also shows information about the start and the end times of the auction as well as the seller's username and his/her rating (in parentheses). The bottom of the page concludes with detailed information about the history of bids. Specifically, starting with the highest bid, the page displays the bidder's user name, his/her rating and the time and date when the bid was placed. Notice that the bid history is arranged in descending order by the bid amount, not by the arrival time. That is, it is quite common that, although a proxy bid is displayed higher up in the bid history, its arrival actually precedes the arrival of some of the bids that appear below it.

2.3 Collected data on Auctions for Palm M515 PDAs

The data for this study were bid histories of 176 closed auctions for *Palm M515* Personal Digital Assistant (PDA) units on eBay.com. We chose the Palm M515 since it proved to be a very popular item, with a multitude of different auctions every day. We collected all the information contained on the bid history pages including bidders' IDs, their ratings, and the time and amount of the bids they placed, as well as information about the start and the end of the auction. All of the auctions were 7 days long and took place from Mid-March through June 2003. Figure 2 displays the bid times and amounts for the 176 auctions (the data are aggregated across auctions). It can be seen that most of the bids arrive on the last day of the auction and that bids range between \$0.10 and \$283.50. The aggregated data are plotted in order to give a general picture. However, the methods that we describe in the next sections treat each auction individually.

3 Data Smoothing

Let t_{ij} denote time that the ith bid was placed, $i = 1, \ldots, n_j$, in auction j $(j = 1, \ldots, N)$. The number of bids n_j varies from auction to auction. In our data we had $N = 176$ auctions and each auction lasted exactly 7 days. Thus, using a daily scale we have $0 < t_{ij} < 7$.

Let $\tilde{y}_i^{(j)}$ denote the bid amount placed at time t_{ij}. For a single auction j, we treat the $\tilde{y}_i^{(j)}$ values as noisy realizations of a continuous curve, pertinent to the bidding dynamics of the auction. In the following, our goal is to reconstruct this curve by applying appropriate smoothing techniques.

3.1 Data Pre-processing

In Figure 2, which displays $\tilde{y}_i^{(j)}$ vs. t_{ij} for all i and j, we can see that the bids follow an upward-sloping curve with most of the bids arriving in the final moments of the auction. Thus, in order to capture the bidding activity, especially at the end of the auction, we transformed the bids into log-scores. A second data pre-processing step was necessary due to the irregular arrival of the bids. We linearly interpolated the raw data in each auction and subsequently sampled the interpolated data at a common set of time points t_i, $0 \leq t_i \leq 7$, $i = 1, \ldots, n$, therefore we drop the subscript j.

For the jth auction, let $y_i^{(j)} = y^{(j)}(t_i)$ denote the value of the interpolated bid amount sampled at time t_i. We represent each auction by a data vector of equal length

$$\mathbf{y}^{(j)} = (y_1^{(j)}, \ldots, y_n^{(j)}). \tag{1}$$

3.2 Smoothing Splines

For each auction, our goal is to estimate a curve that, while providing a good fit to the data, does not exhibit excessive local variability. Moreover, in order to capture the bidding dynamics of each auction, we also want to estimate several derivatives of each curve. Smoothing splines are often found to be a good choice that can achieve these goals well (e.g. Ramsay & Silverman, 1997).

Let $\tau_1, \ldots, \tau_L$ be a set of L knots. Then, a polynomial spline of order p is given by

$$f(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \cdots + \beta_p t^p + \sum_{l=1}^{L} \beta_{pl}(t - \tau_l)_+^p, \tag{2}$$

where $u_+ = uI_{[u \geq 0]}$ denotes the positive part of the function u (see Ruppert et al., 2003, for an introduction to polynomial splines). Let $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p, \beta_{p1}, \ldots, \beta_{pL})^T$ be the $(p + L + 1) \times 1$ vector of coefficients in (2). The choices of L and p strongly influence the local variability of the function f. Indeed, larger values of L and p will result in a more "wiggly" f, exhibiting larger deviations from a straight line. One can measure the degree of departure from a straight line by defining a roughness penalty

$$\text{PEN}_m(t) = \int \{D^m f(t)\}^2 dt, \tag{3}$$

where $D^m f$, $m = 1, 2, 3, \ldots$, denotes the mth derivative of the function f. For $m = 2$, for instance, $\text{PEN}_2(t)$ yields the integrated squared second derivative of f which is sensitive to the curvature of the function f.

For each of the N auctions, the goal is to find a function $f^{(j)}$ that minimizes the penalized residual sum of squares

$$\text{PENSS}_{\lambda,m}^{(j)} = \sum_{i=1}^{n} \{y_i^{(j)} - f^{(j)}(t_i)\}^2 + \lambda \times \text{PEN}_m^{(j)}(t), \tag{4}$$

where the smoothing parameter λ controls the trade-off between data-fit, as measured by the first quantity on the right hand side of (4), and the variability of the function f, measured by the roughness penalty $\text{PEN}_m^{(j)}$. Using $m = 2$ in (4) leads to the commonly encountered cubic smoothing spline. Heckman & Ramsay (2000) discuss minimization for more general values of m (see also Heckman, 1997). Let $\hat{\boldsymbol{\beta}}_{\lambda,m}^{(j)}$ denote the vector of coefficients pertaining to the minimizer of (4) for auction j, $j = 1, \ldots, N$.

In this work, we used the module *Pspline* developed by Ramsay (1996) for minimizing $\text{PENSS}_{\lambda,m}^{(j)}$. The module is available for the software packages *Splus* and *R* and can be downloaded either from Ramsay's web page directly (`http://ego.psych.mcgill.ca/faculty/ramsay/ramsay.html`) or from *StatLib* (`http://lib.stat.cmu.edu`). This module uses all the (distinct) points t_i as knots, thus the number of knots is $L = n$.

3.3 Knot Selection

The number and the locations of the knots τ_l in (2) typically have a strong influence on the quality of the estimated curve. While a smaller number of knots will result in a straighter curve, it may not capture some of the local features of the data. On the other hand, choosing too many knots can result in data-overfitting. Although the roughness penalty approach tries to protect against overfitting by constraining the influence of the individual knots, a poor selection of knots could mask some of the features present in the data.

We suggest to choose the number and location of knots according to the observed bid arrival distribution. Consider again Figure 2. We can see that most of the bids arrive in the final moments of the auction. Indeed, only about 50% of all bids arrive within the first 6 days of the auction. Of the remaining 50% of the bids, more than half arrive within the last 6 hours of the auction. Thus, in order to allow the smoothing spline enough flexibility to capture the intensive bidding activity at the end of the auction, our selection of knots mirrors the distribution of bid arrivals. Specifically, we place 7 equally spaced knots every 24 hours along the first 6 days of the auction, that is, $\tau_1 = 0, \tau_2 = 1, \tau_3 = 2, \tau_4 = 3, \tau_5 = 4, \tau_6 = 5, \tau_7 = 6$. Then, over the first 18 hours of the final day, we place knots over shorter intervals of 6 hours each, that is, $\tau_8 = 6.25, \tau_9 = 6.5$, and $\tau_{10} = 6.75$. And finally, we divide the last 6 hours of the auction into 4 intervals of 1 1/2 hours each, letting $\tau_{11} = 6.8125, \tau_{12} = 6.8750, \tau_{13} = 6.9375$, and $\tau_{14} = 7.0000$.

3.4 Estimating Curve Derivatives

Our goal is to capture the dynamics of bidding behavior. While the bidding curve $f^{(j)}(t)$ describes the exact *position* of the bid at any time point t, it does not reveal how fast the auction is *moving*. Attributes that we typically associate with a moving object are its *velocity* (or its *speed*) as well as its *acceleration*. Given an object with a certain mass, velocity is proportional to the object's *momentum*, while acceleration is proportional to its *force*. Velocity and acceleration can be computed for each auction via the first and second derivative of $f^{(j)}(t)$, respectively.

While velocity and acceleration are common attributes of a moving object, its *jerk*, or third derivative, is typically lesser known. The jerk is an important measure when evaluating the destructive effect of motion on a mechanism or the discomfort caused to a passenger in a vehicle. Jerk measures the rate of change of acceleration and an unsteady acceleration will often be perceived as uncomfortable by the passenger. In our application, the third derivative of $f^{(j)}(t)$ will allow us to asses the rate at which the acceleration of bids changes during the course of the auction.

Notice that the order m of the roughness penalty $\text{PEN}_m(x)$ in (3) controls the smoothness of the derivatives of f. For instance, while using $m = 2$ controls the curvature of the function f, it only controls the slope of its first derivative, $D^1 f$. Thus, Ramsay & Silverman (1997) recommend using $m = q + 2$, where q denotes the order of the highest derivative that is desired. In our application, we are interested in derivatives of order up to $q = 3$, so we set $m = 5$. Derivatives can also be estimated conveniently using Ramsay's *Pspline* module.

Figure 3 shows the curve of $f^{(j)}(t)$ as well as its first three derivatives for two sample auctions, auction number 48 and 75. Notice that the bidding dynamics are quite different for these two auctions. For instance, while the bid velocity and bid acceleration decrease towards the end in auction 48, both increase in auction 75. This seems to suggest that bidding behavior can be quite heterogeneous between auctions. We will investigate this interesting observation further in the subsequent sections.

4 Curve Clustering

The conclusion of the previous section has shown that bidding dynamics can be quite heterogeneous, as illustrated for the Palm M515 auctions. In this section we describe a method for grouping auctions into clusters of homogeneous bidding behavior.

Let $\mathcal{F} = \{f^{(1)}, \ldots, f^{(N)}\}$ be the set of all bidding curves pertaining to the N auctions. The goal is to partition $\mathcal{F}$ suitably into subsets of curves with homogeneous bidding behavior. That is, we want to find a partition $\{\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_K\} \subset \mathcal{F}$, $\mathcal{F}_k \cap \mathcal{F}_{k'} = \emptyset$ $(k \neq k')$, $\bigcup_{k=1}^{K} \mathcal{F}_k = \mathcal{F}$, such that all elements of the cluster $\mathcal{F}_k, k = 1, \ldots, K$, are as homogeneous as possible, while exhibiting maximum heterogeneity towards the elements of the remaining clusters.

One question that arises immediately is how classical multivariate methods can be applied to a set $\mathcal{F}$ of continuous curves. Indeed, classical clustering methods typically assume that the objects to be partitioned arrive in the form of finite-dimensional vectors. Of course, one could deduce for each curve a corresponding finite-dimensional vector by sampling the curve at a specified grid. However, while a finer grid will lead to a more accurate representation of the original curve, the increased dimensionality will also lead to computational problems. A different approach that has been explored in the literature only recently (see e.g. Abraham et al., 2003) is to cluster the set of spline coefficients rather than the functions themselves. That is, let $\mathcal{B} = \{\hat{\boldsymbol{\beta}}_{\lambda,m}^{(1)}, \ldots, \hat{\boldsymbol{\beta}}_{\lambda,m}^{(N)}\}$ the set of coefficients obtained by minimizing (4). Since each of the N curves is based on the same set of knots and the same smoothing parameters, heterogeneity within $\mathcal{F}$ is captured entirely by the heterogeneity within the set of coefficients $\mathcal{B}$.

Several clustering algorithms exists. One of the most well-known methods is the *K-means* algorithm (Pollard, 1981, 1982; Bock, 1985). A drawback to K-means, however, is its lack of robustness against outliers (e.g. Cuesta-Albertos et al., 1997). In this work we consider a more robust variant of K-means, the *K-medoids* algorithm.

For two vectors of coefficients $\hat{\boldsymbol{\beta}}_{\lambda,m}^{(j)}$ and $\hat{\boldsymbol{\beta}}_{\lambda,m}^{(j')}$, corresponding to auctions j and j', where $j \neq j'$, let $d_{j,j'} = D(\hat{\boldsymbol{\beta}}_{\lambda,m}^{(j)}, \hat{\boldsymbol{\beta}}_{\lambda,m}^{(j')})$ denote a measure of dissimilarity between $\hat{\boldsymbol{\beta}}_{\lambda,m}^{(j)}$ and $\hat{\boldsymbol{\beta}}_{\lambda,m}^{(j')}$. A variety of different dissimilarity measures exist (e.g. Cormack, 1971; Gordon, 1990). By far the most common measure is the Euclidian distance, $D(\hat{\boldsymbol{\beta}}_{\lambda,m}^{(j)}, \hat{\boldsymbol{\beta}}_{\lambda,m}^{(j')}) = \|\hat{\boldsymbol{\beta}}_{\lambda,m}^{(j)}, \hat{\boldsymbol{\beta}}_{\lambda,m}^{(j')}\|$. The K-medoids algorithm (e.g. Kaufman & Rousseeuw, 1987; Hastie et al., 2001) is an iterative procedure whose goal is to find a partition $\{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_K\} \subset \mathcal{B}$ which minimizes the within-cluster dissimilarity

$$\mathrm{W}_K = \sum_{k=1}^{K} \sum_{j,j' \in \mathcal{I}_k} d_{j,j'}, \tag{5}$$

where $\mathcal{I}_k$ denotes the set of indices pertaining to the elements of the kth cluster $\mathcal{B}_k$, $k = 1, \ldots, K$.

The K-medoids algorithm achieves this goal in iterative fashion, by alternating between two steps. In the first step, cluster centers are determined. That is, given a current data-partition $\{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_K\}$, one finds the observation in the kth cluster $\mathcal{B}_k$ that minimizes the total distance to the other points in that cluster:

$$j_k^* = \text{argmin}_{j \in \mathcal{I}_k} \sum_{j' \in \mathcal{I}_k} d_{j,j'}. \tag{6}$$

Then, $c_k = \hat{\boldsymbol{\beta}}_{\lambda,m}^{(j_k^*)}$, $k = 1, \ldots, K$, is the current estimate of the center of cluster $\mathcal{B}_k$. The second step re-assigns observations to their nearest cluster. That is, given a current set of cluster centers $\{c_1, \ldots, c_K\}$, one finds a new partition by assigning $\hat{\boldsymbol{\beta}}_{\lambda,m}^{(j)}$ to the cluster $\mathcal{B}_k$ for which

$$k = \text{argmin}_{1 \le k \le K} D(\hat{\boldsymbol{\beta}}_{\lambda,m}^{(j)}, c_k). \tag{7}$$

These two steps are repeated until the assignments do not change any further.

A practical issue with K-medoids (as well as with other clustering procedures) is the choice of K^*, the optimal number of clusters. Notice that, for a given set of data, the within-cluster dissimilarity W_K in (5) generally decreases with increasing values of K, with $\mathrm{W}_K = 0$ for $K = N$. Thus, a commonly used heuristic for estimating K^* is to obtain separate solutions for a variety of cluster-numbers, $K \in \{1, 2, \ldots, K_{max}\}$, and examine a plot of W_K as a function of K. While for every additional *relevant* cluster the reduction in W_K will be substantial, the plot will level off if only irrelevant sub-clusters remain. Thus, an estimate of K^* is obtained by identifying a "kink" in the plot of W_K against K (Hastie et al., 2001). Figure 4 shoes the corresponding plot for the 176 Palm M515 auctions. We can see that there is a "kink" at $K = 2$. From this we deduce that there exist two dominant clusters in the data set.

5 Analyzing & Interpreting Bidding Dynamics for the Palm M515 Data

5.1 Analyzing Average Bidding Dynamics

Figure 5 shows the auction curves $f^{(j)}$, $1 \le j \le 176$, their pointwise average as well as pointwise 95% confidence bounds. The first plot shows the curves pertaining to the log-bids and the remaining three plots display the corresponding first three derivatives, respectively.

The plot of the position of the log-bid versus time shows that, as expected, bid values increase as the auction approaches its end. Notice that the average bid does not start at the origin. Rather, it begins at approximately $\exp(4.14) \approx \$62.80$. The reason for this is the minimum bid, which is often larger than zero and which is set by the seller.

The plot of the bid-position also suggests that for the first 5 days or so, the average log-bid increases at an almost linear rate. Indeed, notice that the average bidding-curve for this period of time resembles a straight line. After day 5, however, the bidding-dynamics seem to change drastically. Specifically, the slope of the curve increases steadily from day 5 towards the end, indicating a tremendous amount of bidding activity at the closing of the auction.

Sharp increases in bidding activity towards the auction-end have been observed in several empirical studies on eBay and have been the focus of multiple research efforts. The phenomenon of bids that arrive very late in the auction is often referred to as "bid-sniping." Bajari & Hortacsu (2003), for instance, find that the median winning bid arrives after more than 98% of the auction time has elapsed. Ockenfels & Roth (2001) report that 14% of all bidders submit their final bid in the last five minutes, 9% in the last minute, and more than 2% in the last ten seconds before the closing of the auction. Bapna et al. (2003) who segment bidders based on bidding strategies use the term "opportunists" to describe bidders who place the minimal required bid just before the auction closes.

There are several ways in which a bidder can engage in bid-sniping. Roth & Ockenfels (2002) report that most people perform bid-sniping "by hand," that is, the user stays online until the end of the auction and submits his or her bid manually only moments before the closing. Another way for engaging in late-bidding is through the use of so-called "sniping agents." Sniping agents are either downloadable programs that run on user's computer or web based services to which a user can subscribe. In either form, these applications allow the user to determine the exact time when the final bid should be submitted without watching the live auction.

A commonly found explanation for bid-sniping is that bidders do not want to reveal their private information about a particular item to other bidders in order to avoid bidding wars (Bajari & Hortacsu, 2003). In fact, Ockenfels & Roth (2001) identify several reasons for bidding as late as possible in the auction. Bid-sniping avoids bidding wars with incremental bidders, who, rather than bidding their secret valuation for an item, only place bids that are marginally above the current high-bid in order to gain the auction-lead. Late-bidding also avoids revealing one's true valuation for an item to uninformed bidders who look to others' bids in order to determine their value for the item. While there may be further reasons why people engage in last-minute bidding, a bidder's main concern is typically not to get outbid in the final moments, especially when his or her valuation for the item is higher than the current high-bid. Thus, placing the final bid as late as possible is a viable solution, although it carries the risk of not being recorded by the auction mechanism due to high server and network traffic (e.g. Roth & Ockenfels, 2002).

The underlying bidding dynamics of the Palm M515 become more transparent by looking at the curves of the first three derivatives. These curves reveal that the average increments for the first 5 days of the auction are not exactly linear. In particular, notice that the average velocity, or the first derivative, of the log-bids increases monotonously over the first 1 1/2 days. This suggests a high amount of bidding activity immediately after the opening of the auction. Recall from Figure 2 that a relatively large number of bids arrive very early in the auction. Bajari & Hortacsu (2003) refer to these bids as "early bids". Bapna et al. (2003) use the term "evaluators" to describe early one-time bidders. However, little to nothing is known as to why people submit early bids and how these bids affect the closing price of an item. Bajari & Hortacsu (2003), based on empirical investigations, conclude that early bids do not affect the magnitude of the winning bid. We will revisit early bidding again later on in this section.

Figure 5 also shows that there is a lack of bidding activity between the opening and the end of the auction. In other words, after the arrival of early bids, the bid velocity decreases until after the mid-point of the auction. Notice in particular the negative acceleration (=deceleration) of the log-bids between day 2 and 3. Then, somewhere between day 3 and day 4, deceleration turns into acceleration and the bid velocity increases sharply towards the end of the auction.

Consider again Figure 5. We can see that the average bid jerk (the 3rd derivative) decreases, starting at day 5. In fact, at the end of the auction, the third derivative is almost zero which indicates a near-constant acceleration. A near-constant bid acceleration at the closing of the auction, however, is surprising, especially in light of a tremendous amount of empirical evidence regarding the prevalence of bid-sniping.

A closer inspection of Figure 5 also reveals a relatively large variability among the bidding curves and their derivatives. For instance, while for a large number of the curves the bid acceleration is increasing towards the end of the auction, there also exists a set of auctions for which the bids are decelerating. This observation suggest that, contrary to previous studies, last-minute bidding behavior is not homogeneous across all auctions. Since all of the auctions under study are based on the same item and are collected during the same time period, behavioral differences due to item category variations or changes over time appear unlikely. We investigate this interesting finding further in the next subsection.

5.2 Analyzing Bidding Profiles

In this subsection we discuss the results of curve clustering described in Section 4. After applying curve clustering to the 176 Palm M515 auctions, we identified $K^* = 2$ dominant clusters, as implied by the within-cluster dissimilarity plot in Figure 4. The features of these two auction-clusters are displayed in Figure 6, which shows that auctions from the two different clusters exhibit very diverse bidding dynamics. Specifically, while the average minimum bid for auctions in cluster 1 is only $\exp(3.83) \approx \$46.06$, in cluster 2 it is almost twice as high at $\exp(4.39) \approx \$80.64$. Interestingly, while the minimum bid is lower in cluster 1, the closing price is higher: the average winning bid in cluster 1 is \$236, while it is only \$228 in cluster 2. The bidding dynamics are even more diverse between the start and end of the auction. A glance at Figure 6 reveals two striking observations. First, there appears to be a higher amount of early "bidding activity" in cluster 1. By early "bidding activity", we refer to curves of the average bid velocity, acceleration and jerk over the initial days of the auction. One can see that, over this time-period, curves in cluster 1 demonstrate a stronger sloping and a higher volatility than in cluster 2. Secondly, Figure 6 reveals a higher amount of late "bidding activity" in cluster 2. Indeed, if we now compare the curves during the final moments of the auction, we notice strongly upward-sloping curves in cluster 2, in contrast to the downward-slopes of cluster 1.

While we used the term "bidding activity" rather loosely in the previous paragraph in order to express a general impression gathered from a first inspection of the plots, it actually refers to the information about the "rate of change" that is truly being conveyed in Figure 6. Specifically, while the bid acceleration reveals the rate at which the bids change, the third derivative discloses information about the rate of change of the bid velocity. In order to investigate the interrelation between the derivatives, one can plot, say, the second derivative of a curve vs. its first derivative. A plot of this nature is called a *phase plane* plot and it is often found useful for the study of differential equations. Ramsay & Silverman (2002) introduced the usefulness of phase plane plots for the analysis of functional data (see also Ramsay & Ramsey, 2002).

Figure 7 shows two phase plane plots, for the second and the third derivative, respectively. Focusing our attention initially on the first plot, we can see that, for cluster 1, the bid acceleration is high at the beginning of the auction and then decreases over the first day. Acceleration is proportional to force and thus a high acceleration at the beginning causes an increase in momentum, or velocity, over the subsequent time period. Indeed, we can see that bid velocity increases to near

0.2 over the first day. If we recall the identity

$$\frac{\partial}{\partial t} \log f(t) = \frac{\frac{\partial}{\partial t} f(t)}{f(t)}, \tag{8}$$

then a (log-) bid velocity near 0.2 implies that with every single unit increase in t, the value of the bid increases by 20%.

The dynamics change between day 1 and day 2. Somewhere in between these two days, acceleration turns into deceleration which results in a reduction of the bid velocity down to near 0.14 after day 3. Acceleration again changes its sign between day 3 and day 4 and increases until shortly before day 6. During this time-span, bid velocity increases to almost 0.4. While the velocity continues to increase up towards the auction-end, the bid acceleration drops significantly, down to almost zero at the end. Thus, we can identify four major phases of bidding activity over the course of the 7-day auction of auctions in cluster 1: high activity during the first day, followed by low (or no) activity until mid-way through the auction. Bidding activity then picks-up again until day 6, albeit slower than at the outset of the auction. Finally, over the last day of the auction, while the bids continue to increase, the rate of increase drops significantly, down to almost zero.

While the bidding behavior in cluster 2 is relatively similar to cluster 1 at the beginning of the auction, it is strikingly different during its conclusion. Figure 7 shows that we can also identify the first three phases of bidding activity for cluster 2, albeit on a different scale. Specifically, while cluster 2 also exhibits an increased amount of activity over the first day, the bid velocity is only about half that of cluster 1. Moreover, the second phase of low (or no) activity is longer than in cluster 1 and extends until after day 4. And finally, while during phase three bidding activity increases again, it continues to increase past day 6. In fact, between day 6 and day 7, the bid velocity increases sharply from near 0.22 up to 0.52 at the end. Thus, while auctions in cluster 2 are also marked by an early phase of light bidding activity followed by a longer phase of low activity, the most striking difference occurs at the end where in cluster 2 the rate of bid-increase continues to grow, especially over the last day.

Since the second derivative determines the rate of change of the bid amount, it is interesting to study the acceleration further. Consider the second plot in Figure 7, which describes jerk vs. acceleration. We can see that acceleration in cluster 1 can be broken up into three phases: decrease until after day 2, then increase until before day 6 and a subsequent decrease until the conclusion of the auction. This is in contrast to cluster 2 which shows only two phases: a decrease until day 3 followed by an increase until day 7.

Why do these two clusters of auctions exhibit such different bidding dynamics? Tables 1 and 2 list summary statistics of variables obtained from the bid histories. These summary statistics shed more light on the differences in bidding dynamics. Recall that the average minimum bid in cluster 2 is almost twice as high as the one in cluster 1. Bajari & Hortacsu (2003) find that auctions with a lower minimum bid typically attract a larger number of bidders. Indeed, the average number of bidders in cluster 1 is 12.4, compared to 9.6 in cluster 2 (see Table 2). Moreover, during the first day of the auction, an average number of 2.7 bidders places an average number of 3.9 bids in cluster 1. This compares to an average of 2.3 bidders placing an average of 3.3 bids in cluster 2 (see Table 1). While these differences are statistically insignificant, the higher number of bidders and bids could well be related to the higher amount of early bidding activity observed in cluster 1.

Tables 1 and 2 about here

While auctions from the two clusters differ in their early bidding activity, the most striking difference lies in their late bidding behavior. Indeed, we have seen that while the rate of the bid-increases slows over the last day in cluster 1, it accelerates for the same period in cluster 2. It is interesting to note that the maximum bid after 6 1/2 days (that is, 12 hours before the closing of the auction) is quite different for the two clusters. The average maximum bid at that time in cluster 1 is already at \$204. This compares to only \$179 in cluster 2 (see Table 2). Moreover, the retail value for the Palm M515, as posted on the manufacturer's web page, is \$249. Thus, if one also takes into account additional fees for shipping and handling that are typically the responsibility of the buyer, then the incentive to place additional bids during the last 12 hours of the auction is much lower in cluster 1 than in cluster 2. This is also mirrored in the number of bids that are being placed. While, on average, 18.4 bids are being placed during the first 6 1/2 days in cluster 1, this number is only 11.8 for cluster 2. Not surprisingly, the relationship is reversed over the last 12 hours. In fact, while only 6.1 bids are being placed (on average) over the last 12 hours in cluster 1, this number is 8.2 for cluster 2.

We want to emphasize that, for the variables in Tables 1 and 2, the difference between cluster 1 and cluster 2 is statistically insignificant. In fact, it is interesting to note that for all of the variables in column 1 of Table 2, cluster 1 differs from cluster 2 by approximately one standard deviation. While these summary statistics tie our findings to previously observed phenomena of the online auction literature, more investigation is necessary to fully understand the difference between different types of bidding behavior. This may well serve as a starting point for future research.

6 Conclusions

In this paper, we hope to expose online auctions to the statistical research community. Our goal is manifold. We aim to introduce an interesting and very rich source of data to the statistical community. We also intend to suggest and demonstrate the use of modern statistical methodology for handling and analyzing such data. Thus, we hope to wed the online auction research with statistical methodology and statistical thinking.

The first part of the paper describes the source of online auction data and its generating mechanism. We describe the notion of web agents and discuss their usefulness and power for collecting data from the web. The popularity of electronic commerce has created an enormous amount of data that are instantly available over the internet. Moreover, the development of intelligent agents (e.g. Murch & Johnson, 1998) has made it possible to "collect" these data conveniently and rapidly from one's office chair. Thus, the combination of readily available electronic data sources with new data-gathering mechanisms has opened the gates for fruitful and exciting pastures of new statistical research.

The second part of the paper describes data analysis tools that can be applied to online auction data. Our focus in this work is on the "bid history," that is, the temporal sequence of bids placed during the course of an auction. We apply tools from functional data analysis by fitting a smooth curve to each bid history. We not only visualize the bidding process, but we also capture the bidding dynamics by estimating velocity, acceleration and jerk for each auction. We use curve-clustering methods to derive profiles of different bidding behavior.

In the last part of our paper, we discuss the implications of our analysis and how the results tie in with existing literature on online auctions. We relate the derived bidding profiles with a variety of online phenomena such as early bidding and "bid sniping" (last-minute bidding). The set of tools that can be applied to this type of data is clearly vast and goes beyond the methods described in this work. Thus, we hope to spark ideas and to create further interest for statistical research in this exciting field of data analysis.

References

Abraham, C., Cornillion, P. A., Matzner-Lober, E. & Molinari, N. (2003). Unsupervised curve-clustering using b-spline. *Scandinavian Journal of Statistics* **30**, 581–595.

Bajari, P. & Hortacsu, A. (2003). Winner's curse, reserve prices and endogenous entry: Empirical insights from ebay. *RAND Journal of Economics* **34**, 329–355.

Bapna, R., Goes, P. & Gupta, A. (2003). Analysis and design of business-to-consumer online auctions. *Management Science* **49**, 85–101.

Besse, P. C., Cardot, H. & Stephenson, D. B. (2000). Autoregressive forecasting of some functional climatic variations. *Scandinavian Journal of Statistics* **27**, 673–687.

Bock, H. H. (1985). On some significance tests in cluster analysis. *Journal of Classification* **2**, 77–108.

Cormack, R. M. (1971). A review of classification. *Journal of the Royal Statistical Society A* **134**, 321–367.

Cuesta-Albertos, J. A., Gordaliza, A. & Matran, C. (1997). Trimmed k-means: An attempt to robustify quantizers. *The Annals of Statistics* **25**, 553–576.

Cuevas, A., Febrero, M. & Fraiman, R. (2002). Linear functional regression: The case of fixed design and functional response. *The Canadian Journal of Statistics* **30**, 285–300.

Elyakime, B., Laffont, J.-J., Loisel, P. & Vuong, Q. (1997). Auctioning and bargaining: An econometric study of timber auctions with secret reservation prices. *Journal of Business and Economic Statistics* **15**, 209–220.

Fan, J. & Lin, S.-K. (1998). Test of significance when data are curves. *Journal of the American Statistical Association* **93**, 1007–1021.

Faraway, J. J. (1997). Regression analysis for a functional response. *Technometrics* **39**, 254–261.

Fraiman, R. & Muniz, G. (2001). Trimmed means for functional data. *Test* **10**, 419–440.

Gordon, A. D. (1990). Constructing dissimilarity measures. *Journal of Classification* **7**, 257–269.

Guo, W. (2002). Inference in smoothing spline analysis of variance. *Journal of the Royal Statistical Society B* **64**, 887–898.

HALL, P., POSKITT, D. S. & PRESNELL, B. (2001). A functional data-analytic approach to signal discrimination. *Technometrics* **43**, 1–9.

HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J. (2001). *The Elements of Statistical Learning*. Springer-Verlag, New York.

HE, G. Z., MULLER, H. G. & WANG, J. I. (2003). Functional canonical analysis for square integrable stochastc processes. *Journal of Multivariate Analysis* **85**, 54–77.

HECKMAN, N. (1997). The theory and application of penalized least squares methods or reproducing kernel hilbert spaces made easy. Tech. rep., University of British Columbia, Department of Statistics.

HECKMAN, N. E. & RAMSAY, J. O. (2000). Penalized regression with model-based penalties. *Canadian Journal of Statistics* **28**, 241–258.

HENDRICKS, K. & PAARSCH, H. J. (1995). A survey of recent empirical work concerning auctions. *Canadian Journal of Economics* **28**, 403–426.

JAMES, G. M. (2002). Generalized linear models with functional predictors. *Journal of the Royal Statistical Society B* **64**, 411–432.

JAMES, G. M. & HASTIE, T. J. (2001). Functional linear discriminant analysis for irregularly sampled curves. *Journal of the Royal Statistical Society, Series B, Methodological* **63**, 533–550.

JAMES, G. M., HASTIE, T. J. & SUGAR, C. A. (2000). Principal component models for sparse functional data. *Biometrika* **87**, 587–602.

JAMES, G. M. & SUGAR, C. A. (2003). Clustering sparsely sampled functional data. *Journal of the American Statistical Association* **98**, 397–408.

KAUFFMAN, R. J. & WOOD, C. A. (2003). Running up the bid: detecting, predicting and preventing reserve price shilling in online auctions. Tech. rep., University of Notre Dame.

KAUFMAN, L. & ROUSSEEUW, P. J. (1987). Clustering by means of medoids. In *Statistical Data Analysis Based on the L_1-norm and Related Methods*.

KLEIN, S. & O'KEEFE, R. M. (1999). The impact of the web on auctions: some empirical evidence and theoretical considerations. *International Journal of Electronic Commerce* **3**, 7–20.

KLEMPERER, P. (1999). Auction theory: a guide to the literature. *Journal of Economic Surveys* **13**, 227–286.

KRISHNA, V. (2002). *Auction Theory*. Academic Press, San Diego.

LI, T., PERRIGNE, I. & VUONG, Q. (2003). Semiparametric estimation of the optimal reserve price in first-price auctions. *Journal of Business and Economic Statistics* **21**, 53–64.

LUCKING-REILEY, D. (1999). Using filed experiements to test equivalence between auction formats: Magic on the internet. *American Economic Review* **89**, 1063–1080.

LUCKING-REILEY, D. (2000). Auctions on the internet: What's being auctioned and how? *Journal of inductrial economics* **48**, 227–252.

MILGROM, P. & WEBER, R. (1982). A theory of auctions and competetive bidding. *Econometrica* **50**, 1089–1122.

MURCH, R. & JOHNSON, T. (1998). *Intelligent Software Agents*. Prentice Hall, New Jersey.

OCANA, F. A., AGUILERA, A. M. & VALDERRAMA, M. J. (1999). Functional principal components analysis by choice of norm. *Journal of Multivariate Analysis* **71**, 262–276.

OCKENFELS, A. & ROTH, A. E. (2001). The timing of bidding in internet auctions: Market design, bidder behavior and artificial agents. *AI Magazine* , 79–88.

OGDEN, R. T., MILLER, C. E., TAKEZAWA, K. & NINOMIYA, S. (2002). Functional regression in crop lodging assessment with digital images. *Journal of Agricultural, Biological, and Environmental Statistics* **7**, 389–402.

PELTO, C. R. (1971). The statistical structure of bidding for oil and mineral rights. *Journal of the American Statistical Association* **66**, 456–460.

PFEIFFER, R. M., BURA, E., SMITH, A. & RUTTER, J. L. (2002). Two approaches to mutation detection based on functional data. *Statistics in Medicine* **21**, 3447–3464.

POLLARD, D. (1981). Strong consistency of k-means clustering. *The Annals of Statistics* **9**, 135–140.

POLLARD, D. (1982). Central limit theorem for k-means clustering. *The Annals of Probability* **10**, 919–926.

RAMSAY, J. O. (1996). Pspline: An *s-plus* module for polynomial spline smoothing. Computer software in the *statlib* archive.

RAMSAY, J. O. (2000). Functional components of variation in handwriting. *Journal of the American Statistical Association* **95**, 9–15.

RAMSAY, J. O. & RAMSEY, J. B. (2002). Functional data analysis of the dynamics of the monthly index of nondurable goods production. *Journal of Econometrics* **107**, 327–344.

RAMSAY, J. O. & SILVERMAN, B. W. (1997). *Functional data analysis*. Springer-Verlag, New York.

RAMSAY, J. O. & SILVERMAN, B. W. (2002). *Applied functional data analysis: methods and case studies*. Springer-Verlag, New York.

RATCLIFFE, S. J., HELLER, G. Z. & LEADER, L. R. (2002a). Functional data analysis with application to periodically stimulated foetal heart rate data. II: Functional logistic regression. *Statistics in Medicine* **21**, 1115–1127.

RATCLIFFE, S. J., LEADER, L. R. & HELLER, G. Z. (2002b). Functional data analysis with application to periodically stimulated foetal heart rate data. I: Functional regression. *Statistics in Medicine* **21**, 1103–1114.

ROSSI, N., WANG, X. & RAMSAY, J. O. (2002). Nonparametric item response function estimates with the EM algorithm. *Journal of Educational and Behavioral Statistics* **27**, 291–317.

ROTH, A. E. & OCKENFELS, A. (2002). 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**, 1093–1103.

RUPPERT, D., WAND, M. P. & CARROLL, R. J. (2003). *Semiparametric Regression*. Cambridge University Press, Cambridge.

TARPEY, T. & KINATEDER, K. K. J. (2003). Clustering functional data. *Journal of Classification* **20**, 93–114.

YU, Y. & LAMBERT, D. (1999). Fitting trees to functional data, with an application to time-of-day patterns. *Journal of Computational and Graphical Statistics* **8**, 749–762.

Table 1: Day-by-Day Auction Summary: For each day of the auction, we computed the number of bids (NBids) on that day, the number of distinct bidders (NBidders) who participated in the auction on that day and the maximum bid (MaxBid) placed during the day. The table displays the mean and standard deviation of these values across auctions from the same cluster.

	Daily Averages (and StdDev) in Cluster 1						
	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
NBids	3.9 (2.7)	3.1 (2.7)	2.9 (2.5)	2.7 (2.0)	3.1 (2.4)	4.5 (3.0)	9.4 (6.0)
NBidders	2.7 (1.2)	1.8 (1.0)	1.6 (0.9)	1.8 (0.9)	1.8 (1.0)	2.6 (1.3)	5.1 (2.6)
MaxBid	68 (48)	78 (44)	97 (58)	113 (55)	136 (55)	170 (46)	236 (17)
	Daily Averages (and StdDev) in Cluster 2						
	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
NBids	3.3 (2.5)	2.8 (2.2)	2.8 (2.9)	2.8 (3.2)	2.6 (2.2)	3.1 (2.7)	9.9 (5.6)
NBidders	2.3 (1.2)	1.8 (1.0)	1.6 (0.9)	1.4 (0.7)	1.5 (0.8)	1.8 (1.0)	5.2 (2.5)
MaxBid	81 (54)	82 (51)	115 (55)	139 (49)	147 (46)	155 (39)	228 (19)

Table 2: Cumulative Auction Summary: For each auction, we computed the number of bids, the number of distinct bidders and the maximum bid over the first 6 1/2 days of the auction (column 1), over the last 12 hours of the auction (column 2) and over the entire auction (column 3). Again, we report the mean and standard deviation of these values across auctions from the same cluster.

	Average (StdDev) in Cluster 1		
	First 6.5 Days	Last 12 Hours	All 7 Days
NBids	18.4 (6.5)	6.1 (4.2)	24.2 (8.5)
NBidders	9.6 (2.9)	3.6 (2.0)	12.4 (3.8)
MaxBid	204 (32)	236 (18)	236 (18)
	Average (StdDev) in Cluster 2		
	First 6.5 Days	Last 12 Hours	All 7 Days
NBids	11.8 (7.4)	8.2 (4.4)	18.4 (10.6)
NBidders	6.0 (3.1)	4.5 (2.1)	9.6 (3.0)
MaxBid	179 (35)	228 (19)	228 (19)

eBay.com Item Bid History Page 1 of 2

home | register | sign in | services | site map | help

Browse | Search | Sell | My eBay | Community

Powered By IBM

Search tips

Search titles **and** descriptions

eBay.com Bid History for
PALM M515 COLOR PDA LIKE NEW HANDHELD (Item # 3041545039)

Currently	**$157.50**	First bid	**$60.00**
Quantity	**1**	# of bids	**19**
Time left	**Auction has ended.**		
Started	Aug-16-03 10:34:26 PDT		
Ends	Aug-21-03 10:34:26 PDT		
Seller (Rating)	**daynathegreat (27)**		

View page with email addresses (Accessible by Seller only) Learn more.

Bidding History (Highest bids first)

User ID	Bid Amount	Date of Bid
moonwolfdesigns (481)	$157.50	Aug-21-03 10:33:20 PDT
rondaroo1 (65)	$155.00	Aug-21-03 10:32:52 PDT
moonwolfdesigns (481)	$151.99	Aug-21-03 10:19:00 PDT
rondaroo1 (65)	$150.00	Aug-21-03 10:32:23 PDT
rondaroo1 (65)	$145.00	Aug-21-03 10:32:11 PDT
rondaroo1 (65)	$140.00	Aug-21-03 09:01:49 PDT
cpumpkinbatman (16)	$125.95	Aug-21-03 10:03:09 PDT
cpumpkinbatman (16)	$120.95	Aug-21-03 10:02:45 PDT
moonwolfdesigns (481)	$115.95	Aug-21-03 08:31:09 PDT
quest3487 (68)	$110.25	Aug-21-03 07:48:01 PDT
moonwolfdesigns (481)	$108.35	Aug-21-03 08:28:58 PDT
moonwolfdesigns (481)	$102.75	Aug-21-03 07:25:57 PDT
quest3487 (68)	$100.25	Aug-21-03 07:19:48 PDT
moonwolfdesigns (481)	$100.00	Aug-21-03 07:25:43 PDT
moonwolfdesigns (481)	$95.00	Aug-21-03 07:25:30 PDT
moonwolfdesigns (481)	$90.00	Aug-21-03 07:25:11 PDT

mhtml:file://C:\WINDOWS\Temporary%20Internet%20Files\Content.IE5\S12ZWP6F\eB... 9/12/2003

Figure 1: Partial bid-history for an eBay auction: Notice that the bids are arranged in descending order by the bid-amount. This order, however, does not reflect the arrival of the bids. Indeed, while *cpumkinbatman* placed his/her last bid of $125.95 at 10:03:09, it was not enough to out-bid the $140.00 placed earlier (at 09:01:49) by *rondarool*. Notice also the user rating (in parentheses) following the user ID.

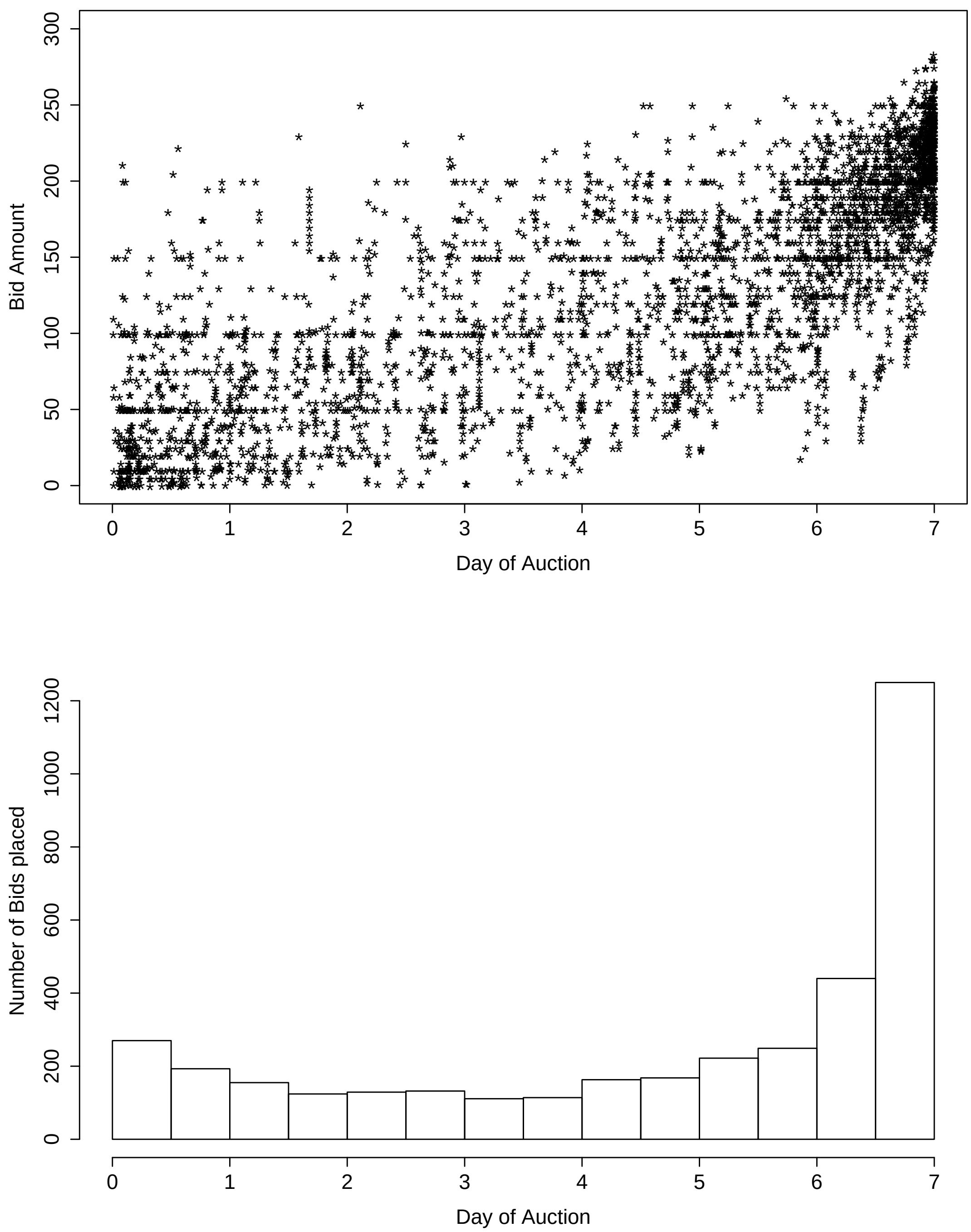


Figure 2: Data for the 176 Palm M515 7-day auctions: The first graph shows the amount of the bid versus the time of the bid, aggregated across all auctions. The second graph shows the distribution of the number of bids over the 7-day auction. Each box corresponds to a 12-hour time interval. Notice that most of the bids arrive in the final moments of the auction.

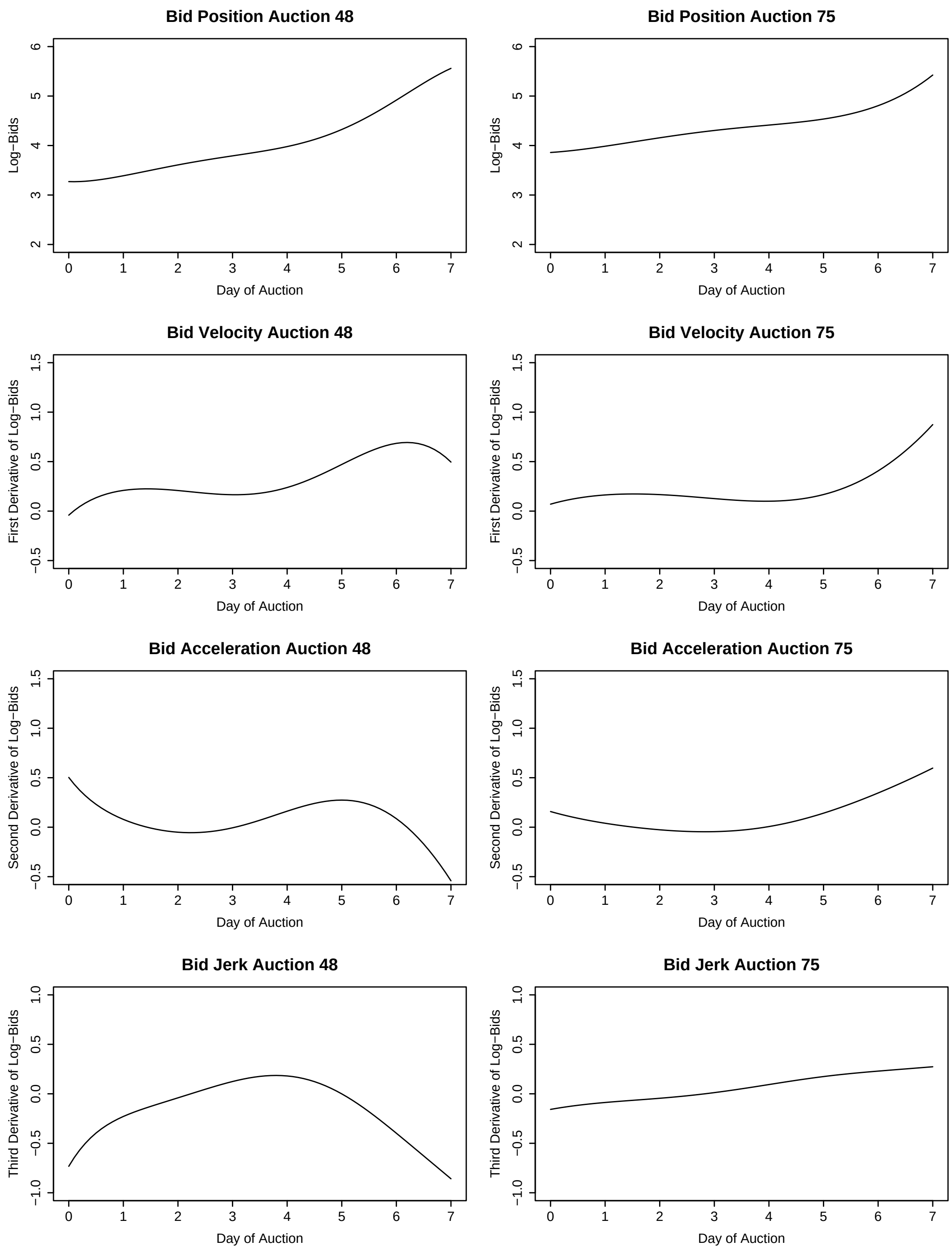


Figure 3: Bidding dynamics for two sample auctions: The left hand side of the graph shows the bid position, velocity, acceleration and jerk for auction number 48. The right hand side corresponds to auction number 75. Notice the different bidding dynamics for these two auctions. In particular, the first three derivatives for auction 48 have a quite shape different compared to auction 75.

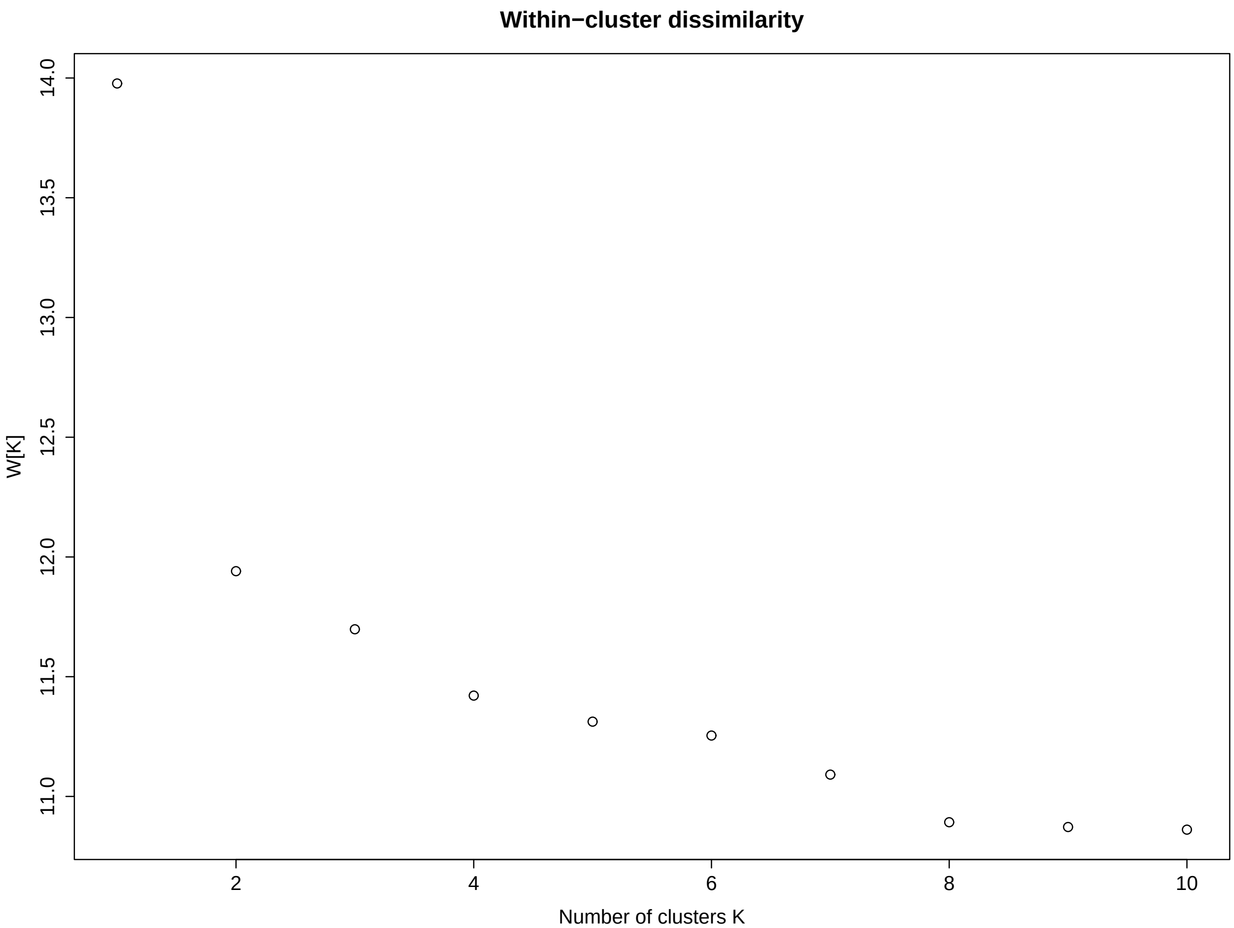


Figure 4: Optimal Number of Clusters: The graph shows the within-cluster dissimilarity W_K in equation (5) as a function of the number of clusters K. Notice the "kink" at $K = 2$, indicating that $K = 2$ is a good approximation to the optimal number of clusters.

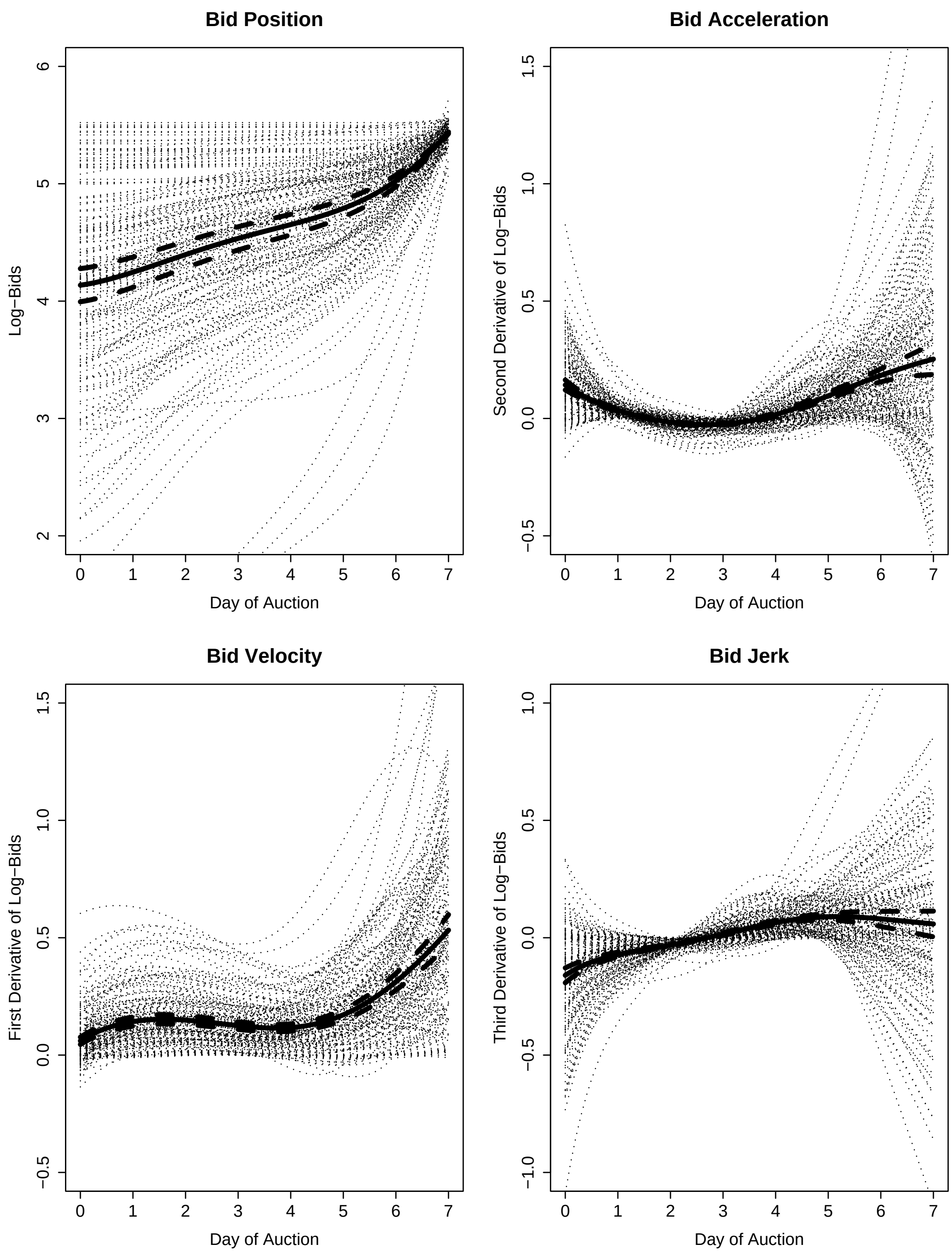


Figure 5: Bidding curves and their derivatives: The thin dashed lines in the first plot correspond to the curves $f^{(j)}$, $1 \leq j \leq 176$, obtained by fitting the penalized spline in equation (2) to each auction. This plot also shows the pointwise average of the curves (thick solid line), as well as their pointwise 95% confidence bounds (thick dashed lines). The remaining three plots are arranged in similar form for the first, second and third derivative, respectively.

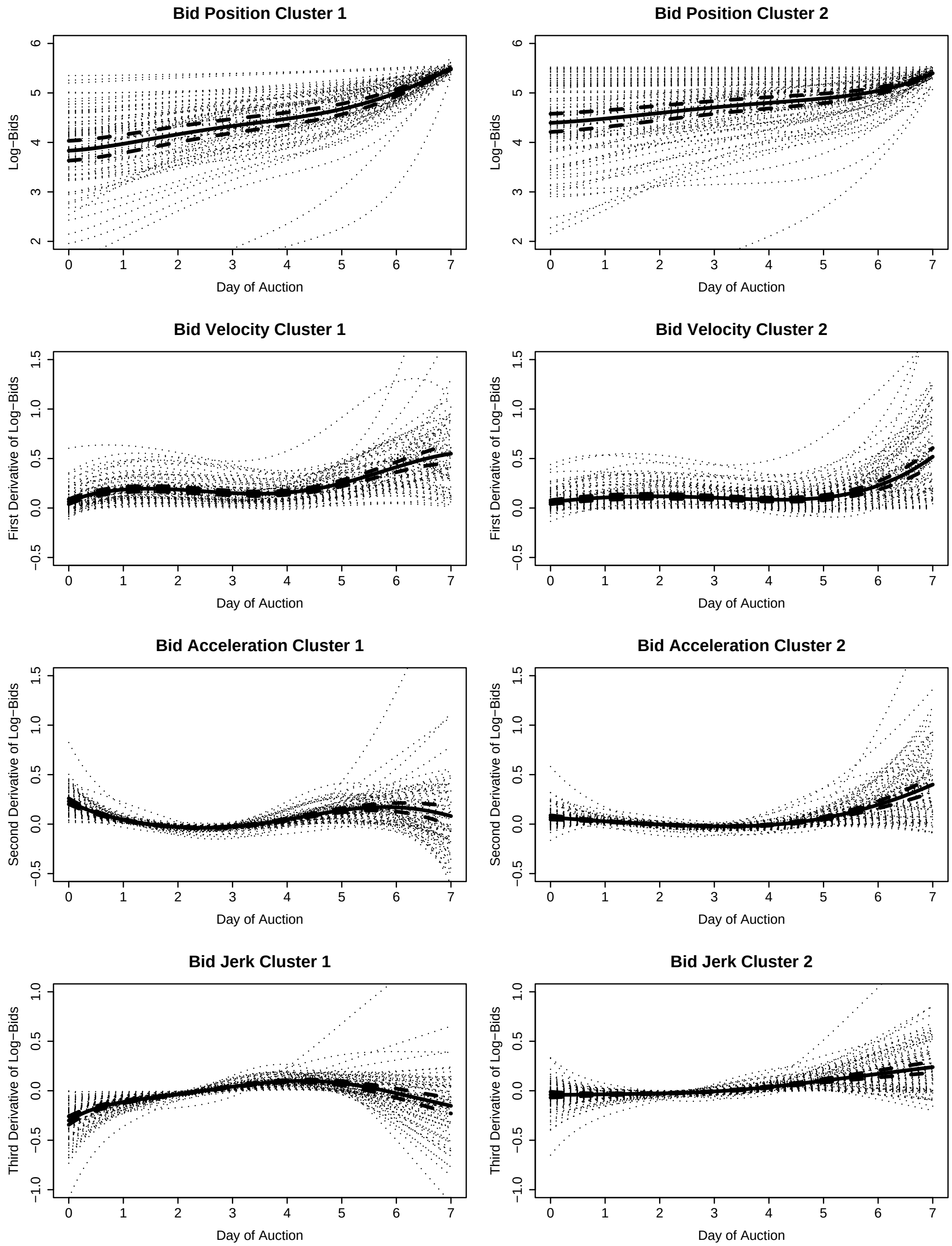
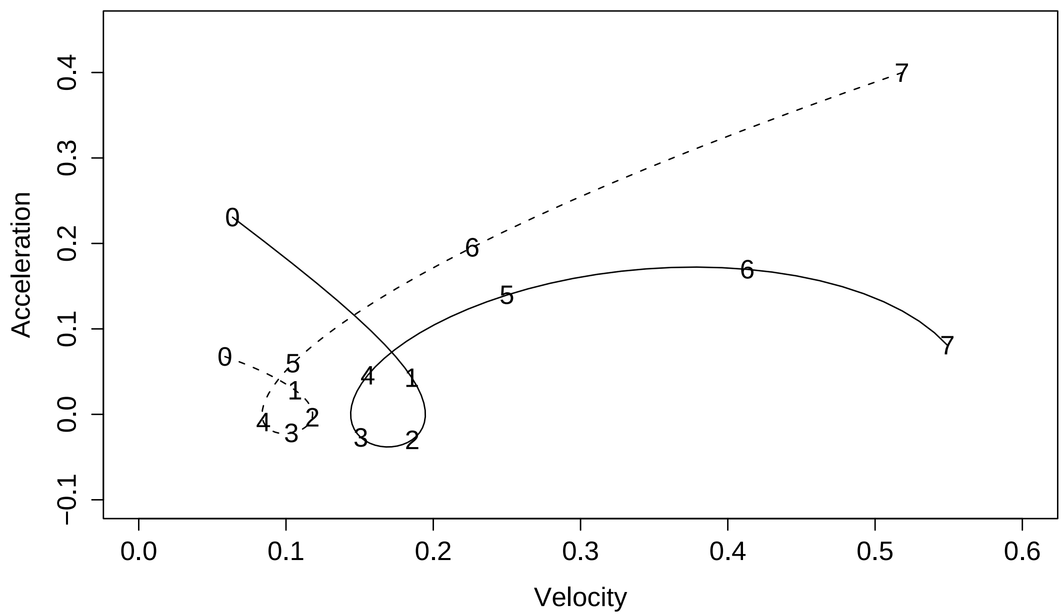


Figure 6: Profiles of two different bidding dynamics: The plots show the assignment of the 176 bidding curves (as well as their derivatives) into 2 different clusters. Notice the different bidding dynamics of the two clusters. Cluster 1 shows high bidding activity at the beginning but decreasing activity towards the end. In contrast, cluster 2 shows rather little activity at the start but increasing activity at closing.



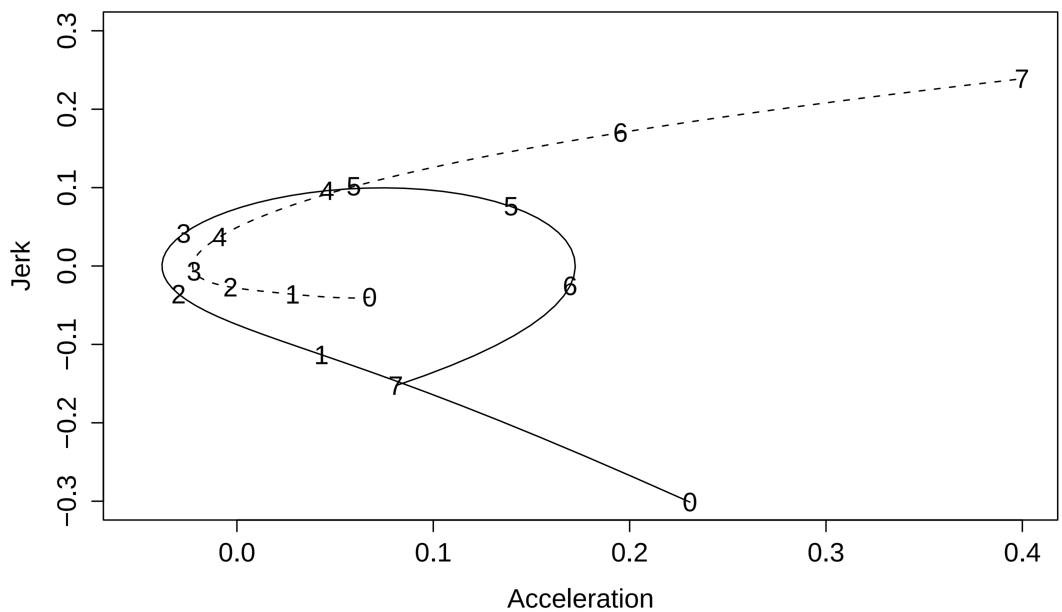


Figure 7: Phase Plane Plots for the Palm M515: The first graph shows the second derivative (acceleration) against the first derivative (velocity). Similarly, the second plot shows the third derivative (jerk) against the second derivative. For each cluster we plotted the average bidding curve. The solid line corresponds to cluster 1, the dashed line denotes cluster 2. The numbers indicate the day of the auction.