ESTIMATING PRICE DYNAMICS IN ONLINE AUCTIONS IN EBAY USING FUNCTIONAL DATA ANALYSIS

Dimitrios I. MADITINOS*
Technological Educational Institute of Kavala Business School
Agios Loukas, 654 04, Kavala, Greece
Tel. +30-2510-462219, Fax. +30-2510-462219
E-mail: dmadi@teikav.edu.gr

Željko ŠEVIĆ
The University of Greenwich Business School
Old Royal Naval College, 30 Park Row, Greenwich,
London SE10 9LS, England, UK
Tel. +44-20-8331-8205, Fax. +44-20-8331-9924
E-mail: Z.Sevic@gre.ac.uk

Athanasios L. IOANNIDIS
Technological Educational Institute of Kavala Business School
Agios Loukas, 654 04, Kavala, Greece
Tel. +30-2510-462219, Fax. +30-2510-462219
E-mail: bigpcstores@yahoo.com

Abstract

Online auctions have become increasingly popular in recent years resulting in online marketplaces and especially eBay having very high annual turnovers. Consequently there is now exposure in the scholarly research on online auctions. In the present study after past literature review and methodologies are presented our methodology and empirical results are summarised. The aim of the present study is to determine online auctions’ price dynamics and the role these dynamics play in the price formation process. The price dynamics considered are the price velocity, price acceleration and price jerk, correspondingly the 1st, 2nd and 3rd derivative of the price’s logarithm. The auction house used in order to extract data was eBay and the exchange item taking part in the present study was Nokia N95. All price dynamics’ curves are developed, presented in figures and their results are discussed. Finally the present paper is a forerunner of studies dealing with the price formation and the forecasting of the final price on online auctions.

Keywords: Online Auctions, EBay, Price Dynamics, Price Formation
1. Introduction

Auctioning can be traced as far back as 500 B.C. (Davidow, 2000). Auctions were always used in order to serve as operational mechanisms so as to assist individuals in exchanging goods. Electronic commerce and more particularly online auctions have become increasingly popular to buyers and sellers alike due to the rapid development in the field of information technology and more specifically of the internet. The whole progress of an auction can help us to extract information about the price dynamics and the price formation.

1.1 Main Concepts

The World Wide Web has dramatically changed how people buy or sell goods nowadays. E-commerce and especially online auction marketplaces have attracted the attention of millions of internet users. People from all over the world can communicate and develop trading relationships (Hahn, 2001). The main advantage of this trading method is the ability that these flea markets have to bring buyers and sellers close in an efficient way, at a great scale and without any time limits.

Our main driver of interest is eBay; the company was founded in 1995 and connects hundreds of millions of people around the world every day. Since the day it was founded, the company’s interest has been to expand in different online business areas. These areas are online marketplaces where eBay has established in total 29 platforms with flagship www.ebay.com. Moreover, in that category eBay finally acquired most of its competitors.

The second area of interest is online payment methods where eBay recently acquired one of the most successful companies related with customer- to-customer payments, PayPal. Finally in the area of communications Skype Company was acquired in October 2005 for 2.6 Billion USD.

In 2003 the company’s net revenues were 21.6 billion dollars; in 2004 there was an increase to 32.7 billion dollars and the same happened in 2005 with net revenue of 45.5 billion dollars. All figures have been taken from the consolidated income statement of eBay Inc. More than 181 million users were registered on eBay at the end of 2005, an increase from the 135 millions of users at the end of 2004. EBay is nowadays the biggest online auction marketplace with several millions of items exchanged every day.

1.2 Main Studies and Findings

The high popularity that online auctions gained over recent years is the reason of the growing interest of scholarly research. Studies on that subject use data from several online auction marketplaces like eBay, Yahoo, uBid and etc. From these online auction houses eBay attracts most of the research because of its high popularity and deep data-accessibility (Jank and Shmueli, 2005). All this data is made possible because of the ability to extract potential data from the web into databases with the help of web crawlers.
Song and Baker (2007) developed an integrated model exploring sellers’ strategies in eBay auctions. Wang, Jank and Shmueli (2004) developed a dynamic forecasting model in order to forecast online auction prices using fundamental data analysis. The important part of their research was that the model had the ability to forecast prices in ongoing auctions in contrast to standard models which many times under-predict the price. Bapna, Jank and Shmueli (2004) studied the price formation and its dynamics in online auctions. Their key finding is that there is little information for the price formation progress in the middle stages of the auction. Moreover, they found that the incremental impact in the rate of the price increase, because of the arrival of an additional bidder, is less at the end of the auction. They also observed that the rate in the price increase is higher in the more expensive items especially at the start of the auction etc. Wang, Jank and Shmueli (2004) studied the dynamic forecasting of online action prices using functional data analysis. They combined static and time-varying information from eBay and developed a very powerful dynamic forecasting system.

The aim of the present study is to determine the auctions’ price dynamics using functional data analysis. Discussion about price dynamics, price velocity and price acceleration will be presented. Price velocity is defined as the speed at which the price is increasing and decreasing during the auction and price acceleration is defined as the rate that this speed changes (Wang, Jank and Shmueli, 2005). In order to catch the price dynamics first it was important to study the bidding process. Bids arrive in spaced time intervals, and more specifically spaced into periods. The two periods with the highest amount of bid arrivals are the start and especially the end of the auction. Moreover, in the present study, some important statistical results have been retrieved taking into consideration nine categorical and five continuous variables. The first categorical variable was the use of the reserve price giving the availability to all sellers to not sell their items under that specific price; the second categorical variable was the condition of the item. The third categorical variable was the auction currency; sellers have the opportunity to sell their items in the currency of their highest interest. In the present study four different currencies were used, the Australian Dollar, the Canadian Dollar, the United Kingdom Pound and the United States Dollar. The fourth categorical variable was the auction duration; eBay gives the possibility to all sellers to use between five auction durations, i.e. 1, 3, 5, 7 and 10 days. The fifth and sixth categorical variables were the possibilities of the seller being a storeowner and the shipping locations the seller send his items to. Finally, the last two were the presence of a subtitle and the auction outcome. After having calculated the percentages of each variable, the average auctions bids were calculated in comparison with the presence of a reserve price. A comparison is also made between the auction currency and the auction locations and between the auction locations and the auction currencies. In addition, the average total bids were compared with the auction duration, the average bids with the store availability of the seller and finally the presence of a subtitle with the average bids set. As stated above, the main object of this study is to determine the price dynamics and this is the object for which everything is developed.

2. Theoretical Background

In the present study price prediction and price dynamics in online auctions and specifically in eBay will be studied. In order to do this, some past findings will be presented here written by specific authors who dealt with the same topic in their articles. Bapna, Wolfgang and Shmueli, (2004) used 1009 eBay auctions and data model in order to study the price
formation procedure and its dynamics. The most important finding of their research dealt with the reality that there is “almost little or no informational content in the middle stages of an eBay online auction’s price formation process”. The above fact strengthens the opinion that “there is significant scope for enhancing the price formation contribution of the early and middle stages of eBay’s auction mechanism” (Bapna, Wolfgang and Shmueli, 2004). It has been found that a further bidder whose entrance in the auction is at a late point of time less influences the velocity in the price enlargement.

It has also been found that the rate of the price increase is higher at the start and at the end of the auction when talking about items with a higher value. Moreover, sellers with high experience, that is to say sellers with many positive feedback points, influence positively the rate and dynamics of the price increase, but this effect is smaller in auctions with a longer duration. Bapna, Wolfgang and Shmueli, (2004) also compared the US and European auctions and observed that US auctions have a 4 per cent higher price in the first half of the auction and subsequent prices are equivalent. In addition, it is concluded that the main differences between US and European auctions occur in the middle of the auction. Although US auctions have a bigger price at the start and the end of the auction duration, price increases in European auctions are faster in the middle part of the auction duration. An additional significant result obtained from the research is that the price level is negatively related to the auction duration when the seller is low-rated. On the other hand, when sellers have a high rating, then their longer duration auctions tend to have a higher price both at the start and at the end of the auction.

Wang, Jank and Shmueli, (2004) observed that in online auctions the uncertainty is reduced more and more as the auction comes to the end. This finding results from the fact that bidders collect more and more information as the auction approaches its close. Fundamental data was extracted from closed auctions and a forecasting model was developed whose results were very near to the real data.

Jank and Shmueli (2005) determined the price dynamics in online auctions using curve clustering. They found that there is a strong relationship between the auction price dynamics and the final price formation. Moreover, in studying the auction price dynamics, they found that price dynamics change sharply during the auction and the changes increase in size especially towards the end of the auction. Shmueli and Jank (2004a) used functional regression analysis in order to study the impact that different factors have (like the opening bid) on the final auction price. Finally Shmueli and Jank (2004b) studied some activities made by bidders on eBay such as “early bidding” and “snipping” and the impact of these activities on the final price formation.

To summarise, it is important to mention that it has been found that if the authors had more attributes for examination such as the presence of a picture of the item in the auction, or the online buying experience of the bidders, the price prediction procedure could be improved further. Finally, it must be added here that because this specific research is made for the data set NOKIA it does not prove that it can be used to predict the final price of online auctions that consist of other kind of products.
3. Methodology

3.1 Introduction

The aim of the present study is to consider the price auction dynamics using fundamental data extracted from an eBay auction selection. In the present chapter the methodology used will be examined and linked with past methodologies. The fact that eBay is a highly popular site, with millions of users worldwide and with a very high yearly turnover, has attracted many researchers interested in the results that could be extracted from the high amount of data given by eBay freely and openly to the public. In the present chapter the most important earlier studies related to our work will be presented attaching significance to the method used in order to extract results. After the presentation of the most important past methodologies, our methodology will be systematically provided.

3.2 Past Methodologies

Wang, Jank and Shmueli (2004) used 185 completed auctions of the item Palm M515 Personal Digital Assistant (PDA) with a duration of 7 days and collected all data about the bid histories. It is observed that most bids came either at the beginning or at the end of the auction. In their study Wang, Jank and Shmueli (2004) focused on the development of a forecasting model in order to estimate the final price for specific items such as the Palm M515 for the same auction duration. The method used in this study to predict final prices is fundamental data analysis. Data is collected and put into context in order to make statistical analysis and subsequently to forecast the final price of the specific item. It is observed that in each auction the price bidding histories develop a price curve. In this study the population used is a large number of curves that are retrieved from the specific auction data. Each curve is capable of describing the heterogeneity of the auction from which the curve’s data is collected. The first step taken by the authors was to estimate all auction curves and use fundamental data analysis. The last method is adopted because of its advantages over the other traditional methods like time series.

To start, the first step taken by the authors using fundamental data analysis was to make some data analysis. This step is very common in the above method. In order to collect the data of the bidding activity with a greater effectiveness, principally at the end of the auction bids have been transformed into log-scores. From the data collected it seems that about 50 per cent of the bids arrive in the last 6 hours of the auction. Consequently, it is important to use “a larger proportion of knots at the end of the auction” because of the fact that this method can easily collect all bids that arrive at the end of the auction. In figure 3.1 below taken from Wang, Jank and Shmueli (2004) the logarithm of the price can be seen and its three derivatives (Bid Velocity, Bid Acceleration and Bid Jerk) in relation with the day of the auction. From the first diagram in figure 1 it is clear that the bids are depicted on an increasing straight line leading to the consideration that the researchers are facing a linear regression problem. The three other plots are showing the price’s dynamics that are velocity, acceleration and jerk.

The first derivative, which is the price velocity, increases at the beginning of the auction until the third day, after that and until the fifth day it declines and finally rises again till the end of the auction. The steady increase in the price velocity in the two final days is a result of the
enormous amount of bidding activity during these days. The big change in the bidding dynamics can also be seen in the graphs of price acceleration, which is the second derivative, and jerk, which demonstrates the third derivative. The main attribute of online auctions is the speedy modification in dynamics. Because of the fact that derivatives play an important role in the change of the price, they use them as forecasting tools. The next step taken by the authors was the development of a model that is capable of estimating and forecasting the auction dynamics.

After the model was developed the sample was divided into two parts. The first part, which made up the 60 per cent of the whole sample, was the training set and the second part 40 per cent was the validation set. The model was put on the training set and the validation set was used in order to measure the model’s forecasting accuracy. Authors used the second model in order to forecast the dynamics on the last day of the auction. Firstly, the parameters are estimated based on the data taken from the training set and with the assumption that the data from the transaction on the final day is not yet giving forecasting accuracy is tested. The accuracy is tested using the data retrieved from the valuation model and the real data taken from eBay. From the comparison, 95 per cent of the “prediction interval is strictly positive”.

Figure 1: Bid Dynamics of One Auction

3.3 Proposed Methodology

Every day millions of auctions and selling items are merchandised on eBay meaning that a large amount of data is publicly available. In the present study, 3-Clicks Excel Add-In was used in order to retrieve all important information. This Excel “plugin” allows all eBay users to obtain eBay auction data directly from Excel. This research will maintain 2 different parts. In the first approach 428 auctions of Nokia N95 will be used that have been manually and randomly collected between 18/07/07 and 20/08/07. With the first part - the opportunity is given to obtain some general statistical information from the auctions that deal with the item Nokia N95 meaning that all different auction durations will be included. In the second part 227 auctions of Nokia N95 having a 7-day duration will be used in order to extract the main and more important results. In the second part a deep analysis will be performed using functional regression analysis. The final price of the auctions will be examined in relation to their primary factor of influence. In order to study the effect that different variables have on
the fast changing price dynamics, functional regression analysis will be used. The method used is comparable to classical regression with the difference that the first operates on functional objects such as curves, shapes and objects and the second gives its answers with the use of vector-valued responses (Wang, Jank and Shmueli, 2005). In our study, continuous curve will be used that will depict the price evolution between the start and the end of the auction as the fundamental object. Moreover, eBay\textsuperscript{' } auctions will be used as data outputs. The results in part 2 will be extracted using auctions with the same duration. From the results of the present study further research can be done using auctions with different auction durations in the same study.

In order to do functional data analysis the first step is to recover as said in (Wang, Jank and Shmueli, 2005) “from the observed data the underlying continuous functional object” and that is done with smoothing techniques. As it is available in (Ramsay and Silverman, 1997) the recovery stage is done through some data pre-processing steps. It is important to mention here that in this study 7 pre-processing steps have taken place. After the pre-processing steps and having the original sample, in order to capture the bidding activity, its bids have been transformed into log-scores and the time of the auction was spaced in equal time intervals. According to the model used by (Wang, Jank and Shmueli, 2005) and the empirical evidence taken from the bid collection on the item Nokia N95 the 7-day duration of the auctions used in the present research will be spaced in 70 “smoothing spline knots” (Wang, Jank and Shmueli, 2005) given by \( Y = \{0.1, 0.2, ..., 6.8, 6.9, 7\} \). The importance of a high number of “smoothing spline knots” is of great significance due to the large amount of bids collected and the increasing demand for accuracy. Having collected all important information about the bid histories, the mathematical formula \( \lim_{x \to x_0} = \frac{f(x)-f(x_0)}{x-x_0} \) was used in order to calculate the first, the second and the third derivative of the price’s logarithm. In this formula \( x \) is the time and more specifically the time intervals used and \( f(x) \) is the price for each bid set. The aim of the present study is to summarise the price dynamics for all auctions in the same graphs respectively. The price dynamics are the price evolution, price velocity, price acceleration and price jerk which will be totally presented in the next chapter where the empirical results will be available.

4. Empirical Results

4.1 Sample Selection

The sample selected in this study is 655 auctions on Nokia N95 mobile phone. From the 428 auctions that will be used in the first part, general statistical results will be obtained. In the second study, 227 auctions will be used that will have a 7-day auction duration. For the second study, 6 eBay sites have been used. It was necessary to use more than one site since it was quite difficult to find 7-day auctions. The 227 auctions used in the second study have been obtained between 20/08/2007 and 30/08/2007 and were used in order to study the price dynamics. For each of the auctions used in this study the bid histories are collected and classified in chronological order. Moreover, all important data that these auctions produce is extracted such as: opening bid, seller feedback score, reserve price, item condition, bidder rating etc. The above information constitutes the variables that will be compared with final price, the dependent variable. The main target of this study is to explore the main factors of the price and how these factors affect the final price development.
4.1.1 Part 1

In part 1 significant conclusions have been extracted taking into consideration 9 categorical variables. Moreover, percentages of each variable are given and each of these variables has been re-marked in order to obtain a better view of its exact meaning. The next step was to relate the average bids with the availability of a reserve price and it is concluded that the variable reserve price helps the auction to gain more bids. Next the relationship between the auction currency and the place where the item is located is demonstrated. As concluded, the auction currency has a strong relationship with the place an item is located. Moreover, the relationship between the auction currency and the availability of shipping locations is studied and it is found that these two variables are strongly related, meaning that for example an auction using the Britain Pound is in a high percentage only in Great Britain. In addition, the auction duration is related to the average bids set and the percentage of failed auctions in each auction duration category. As it is empirically retrieved, the average auction bids grow as the auction duration grows except when the duration is 7-days and 10-days. Moreover, it is studied whether or not the seller of the item is a storeowner in eBay and the presence of a subtitle in relation to the average bids set are given. It is concluded that the use of a subtitle helps the auction to gain more bids. The use of a subtitle attracts the attention of more potential bidders but costs the sellers a fee. In contrast to the general view it is found that auctions belonging to sellers who are not storeowners in eBay collect more bids.

4.1.2 Part 2

The data used in part 2 was 221 7-day duration auctions dealing with the item Nokia N95 mobile phone, after the pre-processing steps, only 104 7-day duration auctions are used which will not break any of the rules set in the pre-processing stage. Beginning in part 2 a summary of the categorical and continuous variables is made. Firstly, as in part 1 the categorical variables are given together with their percentages. Moreover, in part 2 the mean, median, minimum, maximum and standard deviation of 5 continuous variables are extracted. These variables are the start price, the final price, number of bids, seller feedback score and the bidder feedback score. As mentioned in the methodology part the 7-day auction duration is spaced in time intervals. Using this fact the figure 2 below is presenting all bids in relation with the time intervals.

*Figure 2: All 104 Auctions’ Bids, Staring Prices and Time Intervals*
The empirical evidence from figure 2 clearly presents the typical feature of eBay auctions, meaning that most of the bids arrive at the end of the auction. This fact is caused by the snipping method used by buyers. As concluded 51.14 per cent of the bids arrive during the last 12 hours of the auctions. After presenting the bid arrival distribution and obtaining important results about the late bidding process, the price dynamics are studied. Starting with the study of the price dynamics in figure 3, we will use the logarithm of the price in the vertical axis and the time intervals in the horizontal axis. In order to develop figure 3 all price logarithms and time intervals have been summarised in one Microsoft Excel 2007 worksheet. In figure 3 the solid line is a polynomial function with 2nd order, as can be seen the points present the bids set for each time interval. As expected, the price increases monotonically towards the end. The polynomial function presented resembles a straight line; however, the rate of increase does not remain constant. The solid line gives the average price evolution increase for all 104 auctions. Because of the fact that there are many differences in starting prices other auctions start at low prices and shoot up very fast and other auctions start with high prices and remain relatively high throughout the duration of the auction so there are differences in each of their price evolution progresses.

Figure 3: Price Evolution for 104 7-Day Auctions of Nokia N95

In figure 4 the price velocity is presented with the polynomial function and the points are the 1st derivative of the price logarithm in each time interval. From figure 4 it can be seen that the
average price velocity is an almost straight line declining from the first day until the last day of the auction.

*Figure 4: Price Velocity for 104 7-Day Auctions of Nokia N95*

In order to better capture the price dynamics, two more derivatives of the price logarithm are studied here. Before figures 5 and 6 are given presenting the price acceleration and the price, it is important to see table 1 giving the cases of price dynamics availability. The next price dynamic that is examined is price acceleration that shows the change in the price velocity. From figure 5 it can be seen that the polynomial function presenting the price acceleration has some differences from the function of the price velocity. Price acceleration is declining from the start of the auction until day 4 and then starts to increase until the end of the auction. The meaning extracted is that, until the fourth day, bidders set bids whose amount intervals decrease. In contrast, after the fourth day bidders set bids whose amount intervals increase until the end of the auction.

*Table 1: The Cases of Price Dynamics Availability*

<table>
<thead>
<tr>
<th>Cases</th>
<th>1-Bid Auction</th>
<th>2-Bids Auction</th>
<th>3-Bids Auction</th>
<th>4-Bids Auction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price Evolution</td>
<td>Available</td>
<td>Available</td>
<td>Available</td>
<td>Available</td>
</tr>
<tr>
<td>Price Velocity</td>
<td>0</td>
<td>Available</td>
<td>Available</td>
<td>Available</td>
</tr>
<tr>
<td>Price Acceleration</td>
<td>N/A</td>
<td>0</td>
<td>Available</td>
<td>Available</td>
</tr>
<tr>
<td>Price Jerk</td>
<td>N/A</td>
<td>N/A</td>
<td>0</td>
<td>Available</td>
</tr>
</tbody>
</table>

*Figure 5: Price Acceleration for 104 7-Day Auctions of Nokia N95*
The final object examined is the change in the average price acceleration of all 104 auctions. The average price jerk $f'''(x)$ is available in figure 6 and it can be seen that it increases from the start until day 2.5 and then starts to decrease until the end of day 7.

*Figure 6: Price Jerk for 104 7-Day Auctions of Nokia N95*

5. **Concluding Remarks**

The above study, being a forerunner of a study that examines the price formation and the price forecast can be a very useful tool in the hands of the buyer and the seller, but also when used by the auction marketplace. In this study we use 655 auctions of Nokia N95 in order to demonstrate how functional data analysis can be used in order to understand the price dynamics on online auctions in eBay. We believe that the researchers’ understanding of the price formation will be enhanced with the above procedure. A key conclusion extracted is that in the middle of an auction little or no information is available about the price formation procedure. From the above empirical retrieval the necessity of studies like ours is vital.
The sample includes 655 auctions that were used in order to submit a study with two different parts. In Part 1, 428 auctions were used in order to retrieve statistical results such as the percentage of auctions that used the secret reserve price or not, the percentage of new and used items etc. In discussion of these results, it can be concluded that, in general, auctions in which a reserve price is present collect more bids. Also it is found that the location of the item influences to a great extent the use of the currency. The conclusion extracted next is that until the 5 day auctions the average of bids set increases, after that there is a decrease in the 7 day auctions and then again an increase in the 10 day auctions. Table 4.5 also presents the percentage of failed auctions in each case. It is also found that the 10 day auctions fail the most. Moreover it is found that auctions in which the seller has no store in eBay are generally collecting more bids. Finally it is found that the presence of a subtitle, although it has to be paid for, can help sellers gather more bids in their auctions.

In part 2 of our study the functional data modeling was applied to the sample collected including only 7-day auctions in order to produce helpful results for the discussion of the auction’s price dynamics. When the seven pre-processing steps were finished 104 auctions remained that fulfilled all assumptions set. Firstly it is found that the 50.70 per cent of all bids are set on the last day of the auctions. The above fact is a result of the late bidding strategy used by bidders and the phenomenon of “bid snipping”. The results for the price dynamics start with the price evolution and it is found that the price increases monotonically toward the end of the auctions. In order to capture the rate of the price increase, $f'(x)$ and $f''(x)$ have been developed. From figure 4 it can be seen that the average price velocity is an almost straight line declining from the start until the end of the auction. The meaning is that even though the price increases from the start of the auctions their rate of increase declines. Next as it can be seen from figure 5 the average price acceleration declines from the start of the auctions until day 4, and then it starts to increase until the end of the auctions. In the present study we also go one step further to develop also the $f'''(x)$ in order to capture the change in the price acceleration as well. As seen from figure 6 the average price jerk increases from the start of the auctions until day 3.5, from then on it starts declining until the end of the auction. Another key finding extracted is that the price formation is different between similar auctioned items and not homogeneous as expected (Jank and Shmueli, 2005).

The investigation of price dynamics in online auctions analysed in the present study constitutes an optimum forerunner for studies that deliberate with the price formation and the forecasting of the price in online auctions. In that case price dynamics are incorporated into a forecasting model for ongoing auctions (Wang, Jank and Shmueli, 2004). Moreover, using functional regression on the price dynamics, estimated parameter curves can be extracted. In this case the relationship between the price and its formation variables can be studied.

References


