Exploring Bidder Characteristics in Online Auctions: An Application of A Bilinear Mixed Model to Study Overbidders

by

Mayukh Dass

(Under the direction of Lynne Seymour)

Abstract

The recent advancements in network data modeling such as bilinear mixed models have opened doors to many other social researches that were not possible to explore earlier. In this thesis, we demonstrate an application of a bilinear mixed model for a complex human behavior such as *overbidding* in auctions, i.e. placing a bid of higher value than his or her preset valuation of the item. We use an innovative approach of illustrating auction data in the form of a network. The rich network framework allows us to consider bidder interdependence and examine overbidders (termed in the study as *Reactors*) in the presence of bidders who have influenced them to do so (termed here as *Influencers*). Results show that *Reactors* bid on fewer items, but on those with high pre-auction estimated values. They also tend to bid more in the second half of the auction as compared to the first half. Implications for the auction house managers were also presented.

Index words: Bilinear Mixed Models, Overbidding Bidder Characteristics, Online Auctions, Fine Arts, Bayesian Estimation

Exploring Bidder Characteristics in Online Auctions: An Application of A Bilinear Mixed Model to Study Overbidders

by

Mayukh Dass

B.E., Nagpur University, India, 2000

M.S., The University of Georgia, 2003

A Thesis Submitted to the Graduate Faculty

of The University of Georgia in Partial Fulfillment

of the

Requirements for the Degree

Master of Science

Athens, Georgia

2007

Exploring Bidder Characteristics in Online Auctions: An Application of A Bilinear Mixed Model to Study Overbidders

by

Mayukh Dass

Approved:

Major Professor: Lynne Seymour

Committee: Jeongyoun Ahn
Abhyuday Mandal

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2007

Dedication

To Ani

Acknowledgments

Om Sri Sai Ram

My journey through my masters program in Statistics has been very exciting and a great learning experience. This would not have been possible without the support and help of all the faculty members in the Statistics Department. Particularly, I would like to thank Professor Lynne Seymour, whose guidance and suggestions improved this study. My discussions with her during the project have been very energizing and productive. I would also like to thank Professor Abhyuday Mandal and Professor Jeongyoun Ahn for spending their precious time to read through my work and make constructive comments. Their perseverance and support have been vital to the success of this project.

I would also like to thank Dr. Srinivas K. Reddy for his guidance and suggestions during the early stages of the project. I would also like to thank Dr. Umesh Gaur for providing insights from the bidder's perspective.

I would like to thank my parents for supporting me throughout my life. Their love and support helped me during my hard times. Finally, I would like to thank my girlfriend, Ani, for her continuous support and trust in me.

Table of Contents

Page

Acknowledgments v

List of Figures vii

List of Tables viii

Chapter

1 Introduction 1

2 Auction Data and Bidder Network 4

2.1 Modern Indian Art 5

2.2 Bidder Influence Network 5

3 Bilinear Effects Model 10

3.1 General Modeling Process 10

3.2 Modeling the Effect of Third-Order Dependence 12

3.3 Estimation Process 14

3.4 Bidder Covariates Examined 15

4 Results 18

5 Implications and Future Directions 30

Bibliography 32

List of Figures

2.1 Bid History from an Online Auction of Modern Indian Art. 8

2.2 SocioMatrix of Influencers and Reactors 9

4.1 Plots of Marginal Mixing -1 25

4.2 Plots of Marginal Mixing -2 26

4.3 Plots of Marginal Mixing -3 27

4.4 Plots of Posterior Densities-1 28

4.5 Plots of Posterior Densities-2 29

List of Tables

3.1 Evaluation of K-Latent Dimensions 15

4.1 Data Description 18

4.2 Correlation Matrix of the Bidder Covariates 19

4.3 Posteror Means and Quantile-Based 95% Confidence Interval for Influencer and Reactor Level Covariate 22

4.4 Posteror Means and Quantile-Based 95% Confidence Interval for Major Parameter of the Bilinear-Effects Model 23

4.5 Summary of the Findings 24

Chapter 1

Introduction

With recent advancements in network data modeling such as the bilinear mixed model [6], researchers are taking a fresh look at various social science issues that were not possible to explore earlier. Particularly, with the ability to analyze second and third order dependencies, these new models are helpful in determining complex social interactions such as relationships between nations [8], trust between consumers [16] and in our case, inter-bidder influence in auctions. Bidders often influence each other during auctions [1, 5, 11, 15, 19]. This result into competitive arousal among bidders (also known as opponent effect), which leads to *overbidding:* a bidder bids a higher value than his or her preset valuation of the item. In this study, we analyze data from an online auction of fine artworks to determine the characteristics of those who overbid during auctions, using a bilinear effect model.

In auctions of common value items such as fine art, exact valuation of the item is not definite. Bidders have their own value belief based on the amount of information they gather prior to the auction. During an auction, a bidder tends to change his own valuations based on the value signals of other bids [15]. This leads to interdependence among bidder valuations, which is the key to our study.

Traditional modeling techniques such as the general linear model are not adequate to analyze such interdependent auction data. Further, recognizing bidder interdependence raises the issues of second order dependencies (such as common-influencer effects or common-reactor effects) or third order dependencies (such as transitivity and balance) among bidders. This suggests using a modeling framework that is not only capable of incorporating these bidder dependencies, but is also able to include bidder level covariates. A bilinear effect model [6, 8] is a natural choice for this situation.

A bilinear effect model is based on a generalized regression framework and is capable of handling higher order dependencies simultaneously. It builds on the social relations model [20, 22] that are capable of specifying random effects between subsequent bidders. It is capable of simultaneously considering regressor variables; correlation between overbidders that have same influential bidders, between overbidders bidding on the same item, and between overbidders who are also the influencers of one another (reciprocal effect); as well as third order dependencies such as transitivity, clustering and balance.

To facilitate any analysis, first, we need to capture the interdependence among bidders in some form without losing any details. We represent bid history in the form of dyadic relationships between bidders participating in auctions of common lots [4]. At a fundamental level, two sequential bids from bidders i and j in a lot form a dyadic relationship between them. Specifically, we study cases where bidder j overbids in the presence of bidder i . In this study, we term bidder j as a *Reactor* and bidder i as an *Influencer* and $y_{i,j}$ as the level of influence of bidder i over bidder j.

The online art auction data used in this study is exclusive at various levels. First, fine art is unique and heterogeneous in nature as the artworks are different from each other. Second, art items sold in these auctions range from a few thousand to a few millions of dollars. This result into higher stakes both for the auction house managers and the participating bidders and thus, making our study distinctive from other auction studies on low priced items [3, 5, 11]. Third, at present, art connoisseurs are steadily growing in number due to the investment attractiveness of artworks, thus making a strong case to study bidder behavior in this context.

Prior studies on online auctions consider the bids as the manifestation of the bidder's strategy. It is guided by the choice of items the bidder wants to bid on, the timing of bid (bidding early or bidding late) and the amount the bidder is willing to spend [17]. Recent studies [1, 4, 5] show that bidders influence others to overbid with their bid values and their bidding aggressiveness. Kagel et al. [9] and Kagel and Richard [10] suggest that bidders with weaker product information tend to demonstrate the tendency to overbid. Heyman et al. [5] and Ariely and Simonson [1] further added that the competitive arousal between bidders also plays a crucial role in overbidding, along with other psychological factors such as quasi-endowment effect and escalation of commitment. In what type of lot (collection of art items) do bidders overbid? Do they bid more in the first half or in the second half of the auction? Are they selective in the lots on which they bid, or do they bid on many lots? In this research, we attempt to answer these questions. We configure bidding data in a network form $y_{i,j}$ where bidder i plays a crucial role in persuading bidder j to overbid.

There are three primary focuses in this thesis. First, it demonstrates an application of a bilinear effects model for complex human behavior and emphasizes on the importance of new and advanced statistical techniques available to the social science field. Second, it discusses an interesting approach of representing bidding during an auction as a social network. Such a data configuration captures bidder interactions in auctions that are important for our study. Third, it advances the auction literature by investigating auction participants who overbid. From the managerial perspective, these bidders are of great importance to the auction house as their behavior has implications for the lot ordering, setting pre-auction estimates and other auction design characteristics.

The rest of the thesis is setup as follows. First, we describe the auction data of our research and discuss our unique approach of representing it as a social network. Second, we discuss the bilinear effects model and explain how we use it to determine the characteristics of overbidding participants. Third, we present the results. Finally, we discuss the implications of our work and present directions for future research.

Chapter 2

Auction Data and Bidder Network

Online auctions[1] have become a hot research topic in economics, marketing, management and statistics. Particularly, with the availability of detailed bidding data from online auctions, we are now able to investigate sophisticated human characteristics based on their behavior that were not possible earlier. One of the main features of our study is the uniqueness of our online auction data. We have collected the data from an online auction house called *SaffronArt.com.* This auction house sells only Modern Indian Art and has become a prominent distribution channel of that genre in recent years. More specifically, the data come from a three-day auction where 199 art lots (a unique piece of art such as a painting, a drawing or a sculpture) were sold. Unlike eBay auctions, this auction is in a first-price ascending format: The lots are open at a specific date and time, but they are closed sequentially in a group of 20 to 25 lots during that same date. For example, lots 1-25 may close at 9:00am and lots 26-50 will close at 9:30am. Further, to discourage devious online bidder behavior such as *sniping*[2], the auction has a soft closing time since the closing time extends by three minutes whenever a bid comes during the last three minutes of the auction. This time extension continues until no one bids during a span of three minutes.

[1] Laudon and Traver [12] estimate that online auction sales (C2C and B2C) will top $36 billion by 2007. Revenue exceeded $6 billion in 2006 at eBay, the pioneering online auction firm where everything from paperclips to Ferraris get sold. Even traditional auction houses like Christie's (whose annual revenues are expected to top $4 billion in 2006) are adopting the online model. (http://www.iht.com/articles/2006/07/12/news/auction.php) (http://internet.seekingalpha.com/article/25034)

[2] Sniping is a strategic bidding activity where bids are submitted in the last moments of the auction to allow minimal time to other bidders to react to this bid. Such behavior is prominent in eBay auctions as the auction closes promptly at a specific time.

2.1 Modern Indian Art

Modern Indian Art, with over $100 million in auction sales in 2006, is now one of the leading emerging art markets in the world. Although traditional auctions for Modern Indian Art have existed since 1995, it is only since 2000 that the market has exploded, with values realized at auctions growing at a brisk 68.7% annually (coincidentally, this is when *SaffronArt.com*, the source of our data, started its online auctions of Modern Indian Art). In 2006, online auction sales of Modern Indian Art from SaffronArt.com ($36.76 million) had more sales (of Modern Indian Art) than the traditional auction houses like Sotheby's ($35.29 million) and Christie's ($33.08 million). Further, *SaffronArt.com* sold more art items (537) compared to Sotheby's (484) and Christie's (329) in that year[3]. The top ten Indian artists sold 31% of the lots and contributed to 57% of the total value realized at auctions since 1995. Two of these artists are now ranked in the top 100 artists in the world based on their auction sales in 2005. A new set of emerging artists (the new trendsetters, typically born after 1955) have contributed 2% in value and 3% in lots and are becoming increasingly popular, commanding ever higher prices.

2.2 Bidder Influence Network

Second and third order bidder interdependency affects bidder behavior in online auction [4]. Therefore, traditional statistical techniques such as linear models are not suitable for the bidder analysis. To overcome this issue, we formulate a network data structure between the bidders that captures their interdependencies during an auction.

Determining a bidder's valuation change during an auction is also challenging. No prior information is available on the private-value distribution of those participants who overbid. Therefore, it is difficult to follow the changes made by these participants during auctions.

[3]In 2005, online auction sales of Modern Indian Art by SaffronArt.com were $18.06 million, more than that of Sotheby's ($10.49 million) and Christie's ($14.89 million). *SaffronArt.com* also sold more art items (390) compared to Sotheby's (276) and Christie's (248) in 2005.

Operationally, a bidder's limit or reservation price may be defined by the proxy bid, i.e. the maximum value set with the proxy bidding system. Proxy bidding is a commonly available feature in most online auction houses where bidders set a maximum amount they are willing to pay for the auctioned item, and then lets the auction house to place proxy bids on their behalf until that price. Bidders using this facility have a pre-determined value for the item and use it to stay within that value limit [3]. After placing a proxy bid, if a bidder re-enters the bidding process and places a normal (non-proxy) bid that is higher than the earlier proxy bids, then this participant may be overbidding [11]. This process of bidder identification is also conservative in nature, as bidders who set limits and later exceed them without using the proxy bidding system cannot be tracked. Further, using proxy bidding as the operational definition of a participant's limit may not perfectly represent some bidders' maxima, since such a participant may use the system to test different bids or to simply try out the proxy bid system. Ku and his colleagues [11] performed a survey of the bidders and found that most of them use proxies to set their maxima. The respondents used terms such as *maximum personal limit, what we were willing to spend, the most I was willing to bid,* and *by how I valued it* to explain their proxy bids. This research attempts to study specific patterns of overbidders rather than the just their presence. If the proxies were truly arbitrary actions of the bidders, none of the bidder covariates will have significant effect on the occurrence of overbidding. On the other hand, if overbidding is systematic, some definite characteristics of the overbidders will surface.

In order to consider the interdependencies between bidders, proper representation of the bidding data is necessary. The bid history is transformed into an $N \times N$ influence matrix where N denotes the number of bidders participating in the auction. The value $y_{i,j}$ in each cell of the influence matrix indicates the influence of bidder i over bidder j. There are at least four ways $y_{i,j}$ can be measured.

The first approach, which is a very simplistic and conservative view, assigns a value of 1 to $y_{i,j}$, indicating that bidder i (the *Influencer*) has bid prior to a normal bid from bidder j

(the *Reactor*) but after bidder *j's* proxy bids. Consider the bid history shown in Figure 2.1. *Anonymous 3* is the *Reactor* and he is reacting to the bids of *Anonymous 25, Anonymous 47, raccoon and Anonymous 138*. Therefore $y_{i,Anonymous3}$= 1, where i is an element of the set {*Anonymous 25, Anonymous 47, raccoon, Anonymous 138*}.

This approach considers every bidder i to have the same level of influence on bidder j, which may not be true since the *Reactor* may not weight the bids of other participants uniformly. For example, bidder aggressiveness of *influencers* (number of times they placed a bid) may play an important role in the influence process. To capture this, one may assign $y_{i,j}$ to be the total number of times bidder i has placed a bid before bidder *j's* higher non-proxy bid. In Figure 1, the value of $y_{raccoon,Anonymous3}$ and $y_{Anonymous25,Anonymous3}$ will be 2, and all others with positive interaction will be 1.

Another possible indication of level of influence of *Influencers* over a *Reactor* is the order in which these bidders have placed their bids. A participant whose bid is close to that of the *Reactor* may provide more value information to him than the bids which were placed earlier. This interaction can be captured by computing the difference between the bid ranks (order of the bids) of the *Reactor* and his Influencers. For example, we will weight $y_{Anonymous25,Anonymous3}$ more than $y_{racoon,Anonymous3}$ or $y_{Anonymous138,Anonymous3}$ since the last bids of *Anonymous 25* and *Anonymous 3* are sequential.

Building on the same notion, another more definite approach is to compute the difference in time between the *Reactor's* normal bid and bids of other participants placed prior to his but after his proxy bid is calculated. The shorter the time difference, the larger is the influence. A long time difference indicates that the reactor may have used other resources to update his value belief, thus implying weaker influence. Therefore, $y_{i,j}$ is considered as:

$$y_{i,j} = \frac{1}{\sum_{h=1}^{B} \frac{(t_{k,j} - t_{k,i})}{n_{k,i,j}} \times 100} \tag{2.1}$$

where $y_{i,j}$ is the influence of bidder i over bidder j; the summation index h = 1,2, . . . , B is the lot where bidder i has influenced bidder j to change his value belief of the auctioned item, the variables $t_{k,i}$ and $t_{k,j}$ indicate the times at which bidder i and bidder j submitted

110
Jagannath Panda

High-rise- Bee-rise

Signed and dated in English (verso)
2006
Oil, acrylic and fabric on canvas
78 x 156 in (198.1 x 396.2 cm)

$58,140 - 69,770
Rs 2,500,000 - 3,000,000

Next Valid Bid : $ 106,600 (Rs 4,583,800)

Rank	Nick Name	Amount($)	Amount(Rs)	Bid Type	Date & Time (US EST)
Current Highest Bid 1	Anonymous 3	99,100	4,261,300		Mar 8 2007 10:47:06 AM
2	Anonymous 25	91,600	3,938,800		Mar 8 2007 10:21:57 AM
3	racoon	84,100	3,616,300		Mar 8 2007 6:19:29 AM
4	Anonymous 138	76,600	3,293,800		Mar 8 2007 12:54:47 AM
5	racoon	71,600	3,078,800		Mar 7 2007 8:57:15 AM
6	Anonymous 47	66,600	2,863,800		Mar 7 2007 1:39:24 AM
7	Anonymous 25	61,600	2,648,800		Mar 6 2007 11:51:47 PM
8	Anonymous 3	56,600	2,433,800	Proxy	Mar 6 2007 11:10:21 PM
9	Anonymous 25	51,600	2,218,800		Mar 6 2007 11:10:21 PM
10	Anonymous 3	49,100	2,111,300	Proxy	Mar 6 2007 10:30:00 PM
	Start price	46,600	2,003,800		Mar 6 2007 10:30:00 PM

Figure 2.1: Bid History from an Online Auction of Modern Indian Art.

a bid on lot h, respectively; and $n_{k,i,j}$ refers to the total number of times bidder i has placed a bid on lot h before bidder j changes his value on lot h.

Taking a conservative stand, this study uses the second approach to identify the effect of bidder i over bidder j who is overbidding. Hence we label the overbidding participant (bidder j) as a *Reactor* to the influence of bidder i, termed an *Influencer*. Figure 2.2 illustrates the underlying network structure with *Reactors* and *Influencers* as nodes. Here $x_{1,2}$ indicates the number of times bidder 1 (an *Influencers*) has placed a bid between the proxy and normal bid of the bidder 2 (a *Reactor*).

Reactors $\longrightarrow$

Influencer	Bidder 1	Bidder 2	Bidder 3	.... Bidder 80
Bidder 1	0	x_{12}	x_{13}	
Bidder 2	x_{21}	0	x_{23}	
Bidder 3 . . Bidder 80	x_{31}	x_{32}	0	

Figure 2.2: SocioMatrix of Influencers and Reactors

Chapter 3

Bilinear Effects Model

This study assumes that the characteristics of *Influencers* and *Reactors* are homogeneous within the bidder types. Since the bidder influence data is represented in the form of a socio-matrix $Y = [y_{i,j}]$, there are certain modeling constraints that need to be addressed. First, the model needs to accommodate bidder level covariates. Second, the model should allow second order dependencies (such as reciprocity and common bidder effect in auctions where two interdependent bidders have common *Influencers* or common *Reactors*) and third order dependencies (such as transitivity where the influential effect of a bidder i over bidder j is dependent on the third bidder and balance, where the influential effect of bidder i over bidder j is similar to that of other bidders over bidder j)[1]. Such higher-order patterns of dependence are capable of providing useful information for predictive inference [8, 16]. To facilitate all the above requirements, a bilinear random-effect model with Bayesian estimation is used.

3.1 General Modeling Process

This modeling approach is based on the earlier works that specify and analyze random effects for the originator (*Influencer*) and the recipient (*Reactors*) in a social relations setting. It starts with the description of a simple linear model and builds the complexities around it sequentially. Consider modeling the dyadic influence data with a linear regression model of the following form:

$$y_{i,j} = \beta' x_{i,j} + \epsilon_{i,j} \tag{3.1}$$

[1]Please see Hoff [6] and Hoff and Ward [8] for more details on second order and third order dependencies between bidders.

where $y_{i,j}$ represents the influence of bidder i over bidder j, and $x_{i,j}$ represent the dyadic level covariates. Since we also want to take into account various second and third order dependencies, we assume the error component to have a covariance structure that is exchangeable under identical permutations of the indices *i,j* of the *Influencers* and *Reactors*, respectively. In our case, the influence of bidder i over bidder j is distinct from the influence of bidder j over bidder i, i.e. $y_{i,j}$ and *i,j* are different from $y_{j,i}$ and *j,i*, we can represent the joint distribution of the *i,j's* or the residuals in terms of a linear random-effect model

$$\epsilon_{i,j} = a_i + b_j + \gamma_{i,j} \tag{3.2}$$

where a_i represents the effect of an *Influencer*, b_j represents the effect of a *Reactor* and $(a_i, b_j)' \sim$ multivariate normal $[MVN](0, \sum_{ab})$. Thus,

$$\begin{bmatrix} a_i \\ b_j \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ba} & \sigma_b^2 \end{bmatrix} \right)$$

Further, we also consider $(\gamma_{i,j}, \gamma_{j,i})' \sim MVN(0, \sum_\gamma)$. Therefore,

$$\begin{bmatrix} \gamma_{i,j} \\ \gamma_{j,i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\gamma^2 & \rho\sigma_\gamma^2 \\ \rho\sigma_\gamma^2 & \sigma_\gamma^2 \end{bmatrix} \right)$$

This leads to a covariance structure of the error given by

$$\begin{aligned} E(\epsilon_{i,j}^2) &= \sigma_a^2 + \sigma_b^2 + \sigma_\gamma^2 & E(\epsilon_{i,j}\epsilon_{i,k}) &= \sigma_a^2 \\ E(\epsilon_{i,j}\epsilon_{j,i}) &= \rho\sigma_\gamma^2 + 2\sigma_{ab} & E(\epsilon_{i,j}\epsilon_{k,j}) &= \sigma_b^2 \\ E(\epsilon_{i,j}\epsilon_{k,l}) &= 0 & E(\epsilon_{i,j}\epsilon_{k,i}) &= \sigma_{ab}^2 \end{aligned}$$

where i, j, k, and l represent distinct bidders, σ_a^2 represents the variance in the observations due to the presence of common *Influencers*, σ_b^2 represents the variance in the observations due to the presence of common *Reactors* and ρ represents the correlation of observations within a influencer-reactor pair, and serves as a measure of reciprocity or mutuality[2] in the

[2]This model is also known as the "round-robin" model [20, 22].

bidder influence data [6]. Since the goal is to analyze the characteristics of overbidders, only the bidder specific covariates are considered in the model. Thus,

$$\theta_{i,j} = \beta_0 + a_i + b_j + \gamma_{i,j} \tag{3.3}$$

The above equation 3.3 is modeled such that the dyadic data are unconditionally dependent, but conditionally independent; given the random effects of *Influencers* and *Reactors*. Therefore, $E(y_{i,j}|\theta_{i,j}) = g(\theta_{i,j})$,
$p(y_{12}, y_{13}, \ldots, y_{n,n-1}|\theta_{12}, \theta_{13}, \ldots, \theta_{n,n-1}) = \prod_{i \neq j} p(y_{i,j}|\theta_{i,j})$

In our case, a Poisson model with log-link is appropriate, given that we measure the level of influence $y_{i,j}$ as the sequential bid count between bidder i and bidder j. Thus,

$$g(\theta_{i,j}) = \epsilon^{\theta}_{i,j} \tag{3.4}$$

$$p(y_{i,j}|\theta_{i,j}) \sim Poisson(e^{\theta_{i,j}}) \tag{3.5}$$

Therefore, the covariance pattern for the observations is given by:
$cov(y_{i_1,j_i}, y_{i_2,j_2}) = E[cov(y_{i_1,j_1}, y_{i_2,j_2}|\theta_{i_1,j_1}, \theta_{i_2,j_2})] + cov[E(y_{i_1,j_1}|\theta_{i_1,j_1}), E(y_{i_2,j_2}|\theta i_2, j_2)]$ or
$cov(y_{i_1,j_i}, y_{i_2,j_2}) = E[0] + cov[g(\theta_{i_1,j_1}), g(\theta_{i_2,j_2})] \approx cov(\theta_{i_1,j_1}, \theta_{i_2,j_2}) \times \acute{g}(\beta x_{i_1',j_1}) \acute{g}(\beta x_{i_2',j_2})$,

where the pattern for $cov(\theta_{i_1,j_1}, \theta_{i_2,j_2})$ is the same as the $\epsilon_{i,j}$'s in 3.2. Unlike a linear regression, $E(\hat{\beta})$ and $cov(\hat{\beta})$ are also the functions of higher-order dependence along with the first-order and second-order moments [6].

3.2 Modeling the Effect of Third-Order Dependence

Wasserman and Faust [21] suggest the existence of third-order dependence patterns such as transitivity and balance in dyadic data like the bidder influence data. Transitivity describes the interdependence between three bidders i, j and k such that a high likelihood of bidder i

influencing bidder j and bidder j influencing bidder k leads to higher likelihood of bidder i influencing bidder k. Furthermore, in an asymmetric data structure[3] such as the one analyzed here, a combination of three bidders i, j and k is considered balanced if the product of their residuals $\epsilon_{i,j}, \epsilon_{j,k}$ and $\epsilon_{i,k}$ is positive. For example, the bidder relationships are considered balanced if both bidder i and bidder k influence bidder j, or none of them played any influential role on bidder j. To capture the effect of these higher order dependencies in the context of the regression setting, a latent K-dimensional vector z_i for each bidder is constructed and the inner product $z_i^t z_j$ is added to the error model equation 3.2 as suggested by Hoff [6]. Thus,

$$\epsilon_{i,j} = a_i + b_j + \gamma_{i,j} + z_i^t z_j \tag{3.6}$$

Based on the magnitude and direction of the latent vectors, the inner product $z_i' z_j$ will capture different third-order dependence through the expectation of the third order moment $E(\epsilon_{i,j}\epsilon_{j,k}\epsilon_{k,j})^4$. Further, the incorporation of $z_i^t z_j$ into the linear predictor allows additional moments of $\epsilon_{i,j}$. Thus,

$$E(\epsilon_{i,j}^2) = \sigma_a^2 + \sigma_b^2 + \sigma_\gamma^2 + K\sigma_z^4$$
$$E(\epsilon_{i,j}\epsilon_{j,i}) = \rho\sigma_\gamma^2 + 2\sigma_{ab} + K\sigma_z^4$$
$$E(\epsilon_{i,j}\epsilon_{k,i}) = K\sigma_z^6$$

Since the inner-product term is a fixed effect, it can be considered as a reduced-rank interaction term. This is termed as the *bilinear effect* or *multiplicative interaction.*

To include this inner product of the bidder's latent characteristics in the random effect model, equation 3.3 is re-parameterized as

$$\theta_{i,j} = \beta_{inf}' x_{inf,i} + a_i + \beta_{rea}' x_{rea,i} + b_j + \gamma_{i,j} + z_i' z_j \tag{3.7}$$

[3] $y_{i,j}$ is not equal to $y_{j,i}$

[4] For more information on how the values and directions of z_i represent various third-order dependencies, please refer to [6, 8, 16].

where $x_{rea,i}$ are the *Reactor* specific covariates and $x_{inf,i}$ are the *Influencer* specific covariates in the model.

3.3 Estimation Process

Bayesian estimation is performed to estimate the model shown in equation 3.7. A Markov Chain Monte Carlo (MCMC) algorithm is used to sample values of the *Influencer* and *Reactor* specific parameters from their posterior distribution. Like any MCMC process, the estimation of the bilinear model is done by using a three-step process of sampling from the desired target posterior distribution $p(\beta_{inf}, \beta_{rea}, \sum_{ab}, Z, \sigma_z^2, \sum_\gamma |Y)$. The three steps are:

1. Sampling of linear effects in the model:

   (a) Sample $\beta_{inf}, \beta_{rea}|\beta'_{inf}, beta'_{rea}, \sum_{ab}, Z, \theta, \sum_\gamma$

   (b) Sample $\sum_{ab}, \sum_\gamma$ from their full conditionals

2. Sampling of bilinear effects:

   (a) For each bidder $i = 1, 2, \ldots n$, sample $z_i|\{z_j : j \neq i\}, \theta, \beta, s, r, \sum_z, \sum_\gamma$

   (b) Sample $\sum_z$ from its full conditionals

3. Update $\{\theta_{i,j}, \theta_{j,i}\}$ using Metropolis-Hastings step:

   (a) Propose $\begin{pmatrix} \theta^*_{i,j} \\ \theta^*_{j,i} \end{pmatrix} \sim MVN\left(\begin{pmatrix} a_i + b_j + z'_i z_j \\ a_j + b_i + z'_j z_i \end{pmatrix}, \sum_\gamma\right)$

   (b) Accept $\begin{pmatrix} \theta^*_{i,j} \\ \theta^*_{j,i} \end{pmatrix}$ with probability $\frac{p(y_{i,j}|\theta^*_{i,j})p(y_{j,i}|\theta^*_{j,i})}{p(y_{i,j}|\theta_{i,j})p(y_{j,i}|\theta_{j,i})} \wedge 1$

The `‘gbme’` function[5], written in R (Hoff 2005b) is used to perform the above estimation process and to determine the characteristics of auction participants who overbid (*Reactors*). The prior distributions of the parameters are taken as [7]:

[5] http://www.stat.washington.edu/hoff/Code/GBME/

Table 3.1: Evaluation of K-Latent Dimensions

K	log p($Y\|\hat{\beta}, \hat{a}, \hat{b}, \hat{Z}, \hat{\Sigma}_\epsilon$)
0	-399.45
1	-397.14
2	-392.71
3	-394.33
4	-398.98

- $\beta \sim MVN(0, 80 \times I_{3\times3})$
- $\Sigma_{ab} \sim$ inverse Wishart$(I_{2\times2}, 4)$
- $\sigma_u^2, \sigma_v^2 \sim$ iid inverse gamma(1,1), $\sigma_\gamma^2 = (\sigma_u^2 + \sigma_v^2)/4, \rho = (\sigma_u^2 - \sigma_v^2)/(\sigma_u^2 + \sigma_v^2)$

We used K=2 (latent dimensions) in our analysis. Five different values of K, ranging from 0 to 4, were tested with a four-fold cross-validation procedue as described by Hoff in his seminal paper [6]. The predictive performance for all the K values were roughly the same. The biggest improvement in the marginal likelihood criterion was from K=1 to K=2. Therefore, K=2 was selected. Table 3.1 illustrates our findings.

3.4 Bidder Covariates Examined

Influencers and *Reactors* are examined in the context of three types of characteristics: lot characteristics, auction characteristics and bidder behavior characteristics.

3.4.1 Lot Characteristics

Pre-auction estimates of the lots provided by the auction house and the type of lot (paper or non-paper works) are the two lot (product) characteristics examined in this study. Before the auction starts, potential bidders are exposed to various types of information about the lots. Lot information and provenance provided by the auction house (in their printed catalogs and in their websites), and comments and suggestions of the art experts (in personal blogs, art magazines, etc.), provide information on the value of the art items such as the estimated price, artist information, and previous auction price history of similar paintings by the artists. One of the most distinct contributions to price-formation during an art auction are the pre-auction estimates [2, 18]. These estimates indicate the value of the items as suggested by the auction house experts (like curators, art specialists, etc.). Mei and Moses [14] found these value estimates to have high correlation with the final realized prices of the art items. Therefore, higher the pre-auction estimates, greater is the tendency for the item to fetch a higher price, and so higher stakes are associated with it. Such high stakes lead to higher bidder propensity to clarify and justify their bids. Therefore, bidders look for other value signals from different sources and are inclined to change their value belief during the auction of these lots. Thus, we hypothesize that the bidders who act as reactors tend to bid on lots that have high pre-auction estimates.

The media on which the art item is painted plays an important role in its maintenance and longevity. For example, works on paper are typically of low price as compared to that of works on canvas since paper tends to be more fragile than canvas. Therefore, low financial risk is associated with the purchase of paper items and thus, we hypothesize that the *Reactors* will tend to overbid on canvas works.

3.4.2 Artist Characteristics

Artist characteristics such as reputation and previous auction history of the artists play an important role in the valuation of the art items. Established artists are highly reputed and their works are well recognized in the art market. Most of their works have been resold many times in the market and, thus, their value commonly known to all. Therefore, the works of established artists present a low risk purchase opportunity for the bidders [18]. On the other hand, emerging artists are new to the art market and their works are not well known. Further, not enough works of these artists are sold to estimate their values confidently. Therefore, the values of the works of these artists are highly uncertain, making them high-risk purchases. Thus, Reactors are hypothesized to bid in lots created by emerging artists where they will seek for more information from other bidders.

Like artist reputation, historical market information such as the average price per square inch of an artist's lots or the total number of the artist's lots sold in the previous year provide a signal to the bidders about the market value of the artist [18]. If the value realized by the lots of a particular artist in the previous year is low, present market value of the artist will be low and uncertain. Since evaluators are unaware of how the market will react to the works of these artists, they are high-risk purchases for the bidders. Similarly, if few of the artist's lots are sold in the previous year's auctions, less information is available about the artist's present market value. Therefore, it is highly probable for bidders to rely upon other value signals in these lots. Thus, it is hypothesized that *Reactors* will tend to bid on these lots.

3.4.3 Bidding Characteristics

Auction participants who overbid (*Reactors*) are assumed to wait and consult bids of others participants (*Influencers*) to reduce any uncertainty they have about the value of a lot. Thus, we conjecture that *Reactors* will bid more in the second half of the auction than in the first half. Further, these bidders by the virtue of their behavior tend to be selective in the types of lots on which they bid. Their attachment to the auctioned lots is an integral reason for them to update their private value for the lot. Therefore, we hypothesize that they will participate in auctions of fewer lots than other bidders (i.e. *Influencers*) will. We also included a variable to control for the number of lots won by the bidders. This variable is used to determine whether *Reactors* tend to win more due to overbidding than the *Influencers*. In other words, whether overbidding truly translates to the success of the bidders.

Chapter 4

Results

The description of the bid data is presented in Table 4.1. In that particular auction, 199 lots were sold, and 42 bidders were observed to change their value belief for 63 lots during a Modern Indian Art auction held in December 2005. Eighty bidders participated in the auctions of these 63 lots creating 947 bid instances. Thus, an 80×80 bidder matrix is created and only the influences on those forty-two bidders are considered.

Bidder level covariates are used in the model to determine the characteristics of those who overbid. Table 4.2 presents the correlation between the bidder level covariates. Although most of the correlations were moderate, correlation between the total number of lots won by the bidder and the total number of bids placed in the second half of the auction is in the higher side (0.734). This suggests that bidding early in the auction does not guarantee a 'win' and bidding late in the auction is a significant strategy for the winners in the auction.

Table 4.1: Data Description

No. of Lots Sold	199
No. of Bids	3080
Average No. of Bids per Lot[Range]	15.47 [2, 48]
Average Value of the Lots [Range]	$56,282 [$2,850, $1,351,000]
Average First Bid of the Lots	$19,343 [$650. $300,000]
Average No. of Bidders per Lot [Range]	6.35 [2, 14]
Average Time of Bids (Scaled 0-1)	0.4998
Average Time of Entry to the Auction	0.5386
Average Time of Exit to the Auction	0.8397

Table 4.2: Correlation Matrix of the Bidder Covariates

	(x_1)	(x_2)	(x_3)	(x_4)	(x_5)	(x_6)	(x_7)	(x_8)	(x_9)
Total No. of Lots won (x_1)	1								
Low pre-auction estimates of the Lots Bid (x_2)	0.243 0.0299	1							
No. of Lots sold by the Artists in the previous year's auction (x_3)	0.215 0.0559	0.551 $< .0001$	1						
No. of Lots by Emerging Artist (x_4)	0.373 0.0007	-0.183 0.2657	-0.182 0.0980	1					
No. of Lots by Established Artist (x_5)	0.459 $< .0001$	0.584 0.0092	0.655 $< .0001$	0.210 0.1725	1				
No. of Paper Works (x_6)	0.345 0.0017	0.127 0.2331	0.201 0.6055	0.474 $< .0001$	0.399 $< .0001$	1			
No. of Bids Placed in the First Half of the Auction(x_7)	0.285 0.0105	0.107 0.6351	0.158 0.2972	0.359 $< .0001$	0.431 $< .0001$	0.515 $< .0001$	1		
No. of Bids Placed in the Second Half of the Auctions(x_8)	0.734 $< .0001$	0.204 0.2851	-0.089 0.5026	0.291 $< .0001$	0.264 $< .0001$	0.116 $< .0001$	-0.201 0.0968	1	
Total No. of Unique Lots Bid (x_9)	0.528 $< .0001$	0.239 0.4630	0.243 0.3677	0.666 $< .0001$	0.673 $< .0001$	0.693 $< .0001$	0.606 $< .0001$	0.304 $< .0001$	1

A Markov Chain Monte Carlo algorithm was run for the influence data. Each chain was run for 200,000 iterations. However, since a large number of parameters (*Influencer* and *Reactor* specific covariates, and terms capturing higher order dependencies) is analyzed in the model, only every 50th iteration is stored in order to keep the output file to a reasonable size, as suggested by MacEachern and Berliner [13]. Outputs from the first 20,000 iterations were considered burn-in and were not recorded. The posterior means and quantile-based 95% confidence intervals are presented in Table 4.3 and Table 4.4. Log transformation of the covariates were used in the analysis. The plots of marginal mixings are presented in Figures 4.1, 4.2 and 4.3. The plots of the posterior densities are also illustrated in Figures 4.4 and 4.5. The variables order for the results is same as the correlation table. For example, $bs_1 = br_1$ = *Total no. of lots won*, $bs_2 = br_2$ = *Low Pre-Auction Estimates of the Lots Bid* and so on.

The parameter estimates illustrate that the effect of a pre-auction estimate is positive and significant at 95% C.I. for the *Reactors*. This suggests that bidders who typically change their value belief will bid on lots with high estimated values. Positive but non-significant estimates for the number of paper works sought by the bidders indicate that the *Reactors* tend not to bid on paper art items. Further, none of the artist related covariates were found to be significant at 95% C.I. in explaining the behavior of the *Reactors*.

We also examined the bidder characteristics of the *Influencers* and *Reactors* by compared their bidding frequency during each half of the auction. Results show that the coeffient for the number of second half bids for the *Reactors* is positive and significant at 95% C.I., but the coefficient for the first half bid frequency is not significant. This suggests that *Reactors* tend to bid more in the second half of the auction than the first half. We also found these coefficients to be non-significant for the *Influencers*, suggesting that there is no significant difference in the bidding activity of the *Influencers* in the two halves of the auction. We also examined the coefficient for the total number of unique lots bid by the two types of bidder. We found that the coefficient for the *Reactors* is negative and significant at 95% C.I. (coefficient= -12.1310) and the coefficient for *Influencers* is positive and significant at 95%

C.I. (coefficient=2.2790). This suggests that that *Reactors* participate in auctions of fewer lots as compared to the *Influencers*. A summary of the results is presented in Table 4.5.

Table 4.3: Posteror Means and Quantile-Based 95% Confidence Interval for Influencer and Reactor Level Covariate

Covariates	Influencers	Reactors
Total Number of Lots Won	−0.0971	−3.9724
	−0.0405	**−1.2070**
	0.3892	0.3844
Low Pre-Auction Estimates of the Lots Bid	−1.5897	0.6235
	−0.5940	**3.0415**
	0.3151	5.6251
No. of Lots Sold by the Artists in the Previous Year's Auction	−0.6772	−3.4006
	0.4740	**−0.3595**
	1.5471	2.5692
No. of Lots by Emerging Artists	−0.7331	−1.5737
	0.4255	**1.5785**
	1.5731	5.0032
No. of Lots by Established Artists	−0.8061	−3.0646
	0.4035	**1.4460**
	1.6606	5.0095
No. of Paper Works	−1.8352	−0.8961
	−0.7275	**4.4405**
	0.3990	7.9614
No. of Bids Placed in the First Half of the Auction	0.1092	−0.2071
	0.3700	**1.7230**
	0.9595	3.3674
No. of Bids Placed in the Second Half of the Auction	−1.1192	3.7268
	−0.5150	**5.4900**
	0.0920	7.3632
Total Number of Unique Lots Bid	0.1425	−17.8651
	2.2790	**−12.1310**
	4.0658	−5.7152

Table 4.4: Posteror Means and Quantile-Based 95% Confidence Interval for Major Parameter of the Bilinear-Effects Model

Parameters	Posterior Mean and Quantile-Based 95% C.I.
Common Influencer Variance	2.1639
	3.7710
	6.6169
Common Reactor Variance	33.1427
	49.5085
	76.1923
Influencer-Reactor Variance	-7.0140
	$\mathbf{-1.9805}$
	-2.8638
Error Variance	13.4692
	16.5935
	20.0395
Reciprocity	-0.9980
	$\mathbf{-0.9950}$
	-0.9769
Variance of Latent Dimensions	0.0489
	0.1440
	0.4071
Variance of Inner Dimensions	0.0150
	0.0550
	0.3501
Log Likelihood of $Y_{i,j}, Y_{j,i}$ Modeling Effects	$\mathbf{-392.719}$

Table 4.5: Summary of the Findings

COVARIATES	**INFLUENCERS**	**REACTORS**
Total No. of Lots Won	Not significant	Not significant
Low Pre-Auction Estimates of the Lots Bid	The posterior mean indicates a a positive, non-significant effect suggesting that *Influencers* tend to bid on both low and high estimated items	The posterior mean indicates a positive, significant effect suggesting that *Reactors* tend to bid on lots with high pre-auction estimates.
No. of Paper Works	Not significant	Not significant
Artist's Reputation	Not significant	Not significant
Artist's Price History	Not significant	Not significant
Bid Frequency at Each Half of the Auction	The posterior mean shows a positive but non-significant effect of the first half bid freq. and negative and non-significant effect of the second half bid freq.This indicates that there is no significant difference in bidding frequency of the *Influencers* on both halves of the auction	The posterior mean shows a positive and significant effect of second half bid freq. and negative and non-significant effect of the first half bid freq. This indicatesthat the *Reactors* tend to bid more in the second half as compared to the first half of the auction.
Total No. of Unique Lots Bid	The posterior mean indicates a a positive and significant effect. suggesting that *Influencers* tend to bid on large number of lots.	The posterior mean indicates a negative, significant effect suggesting that *Reactors* tend to bid on small number of lots.

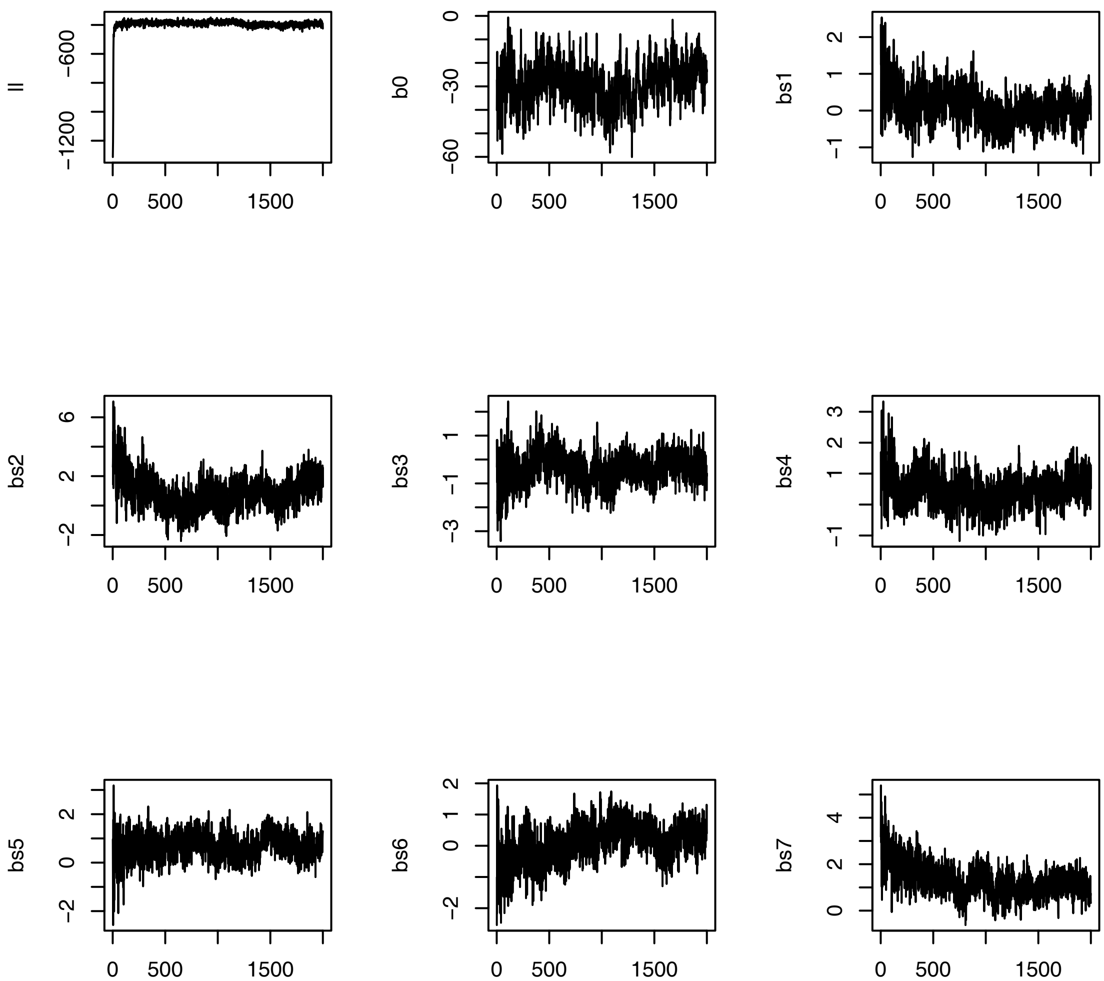

Figure 4.1: Plots of Marginal Mixing -1

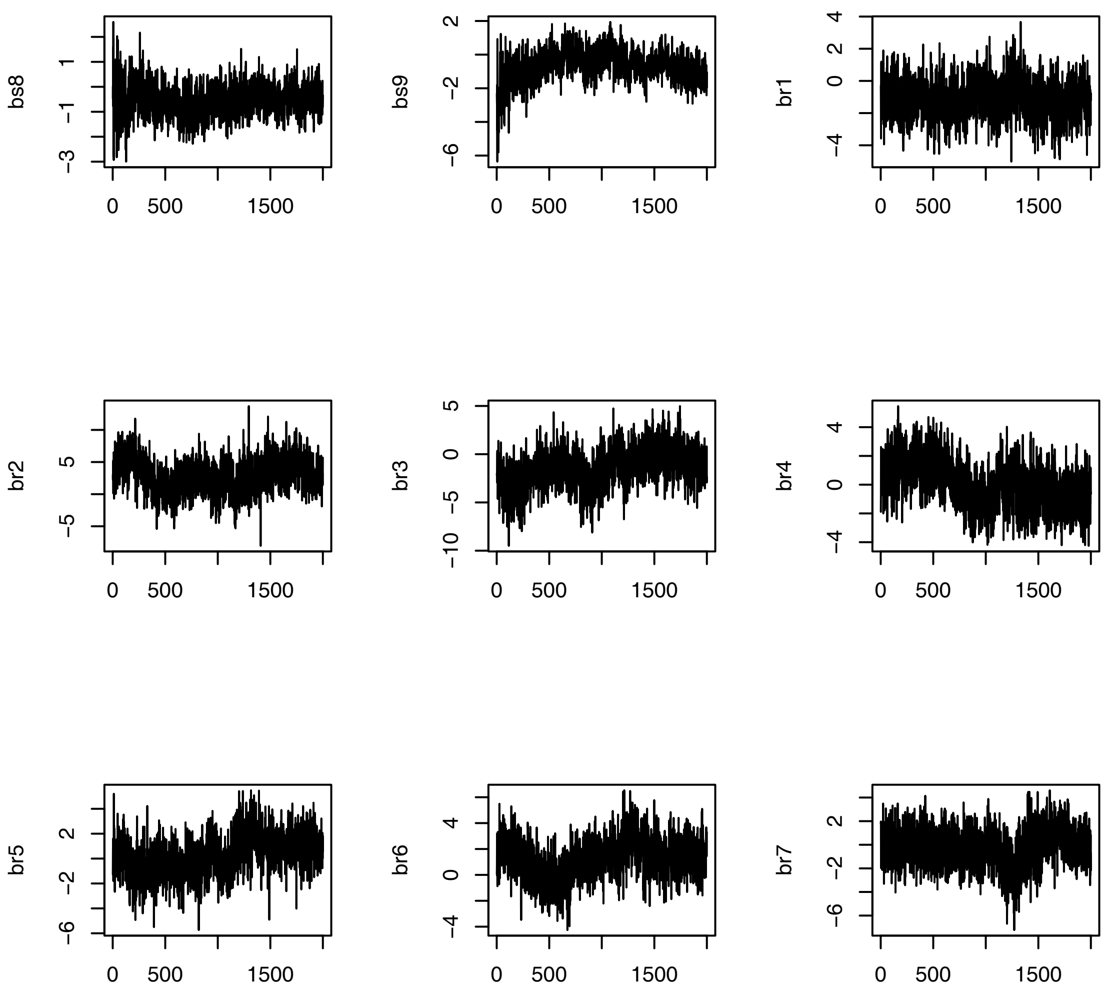

Figure 4.2: Plots of Marginal Mixing -2

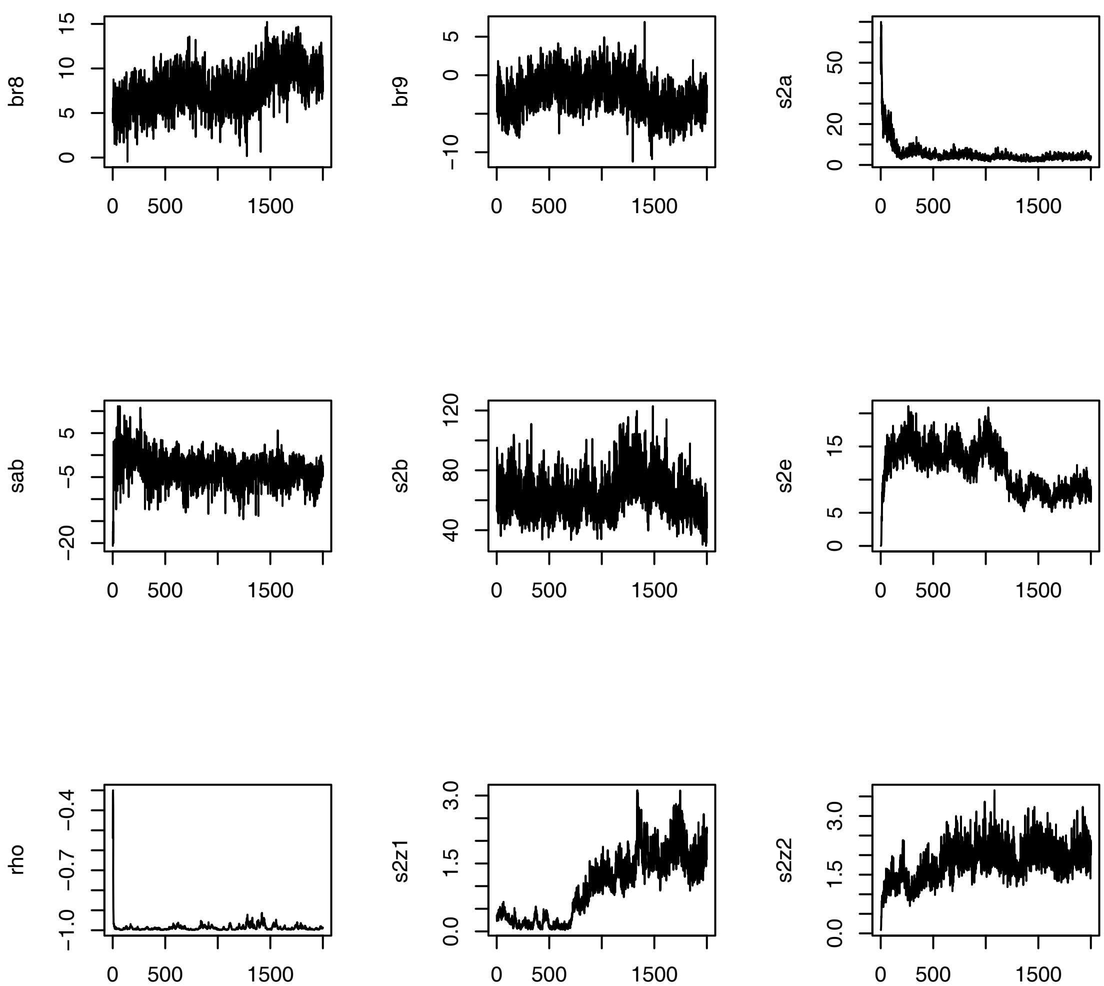

Figure 4.3: Plots of Marginal Mixing -3

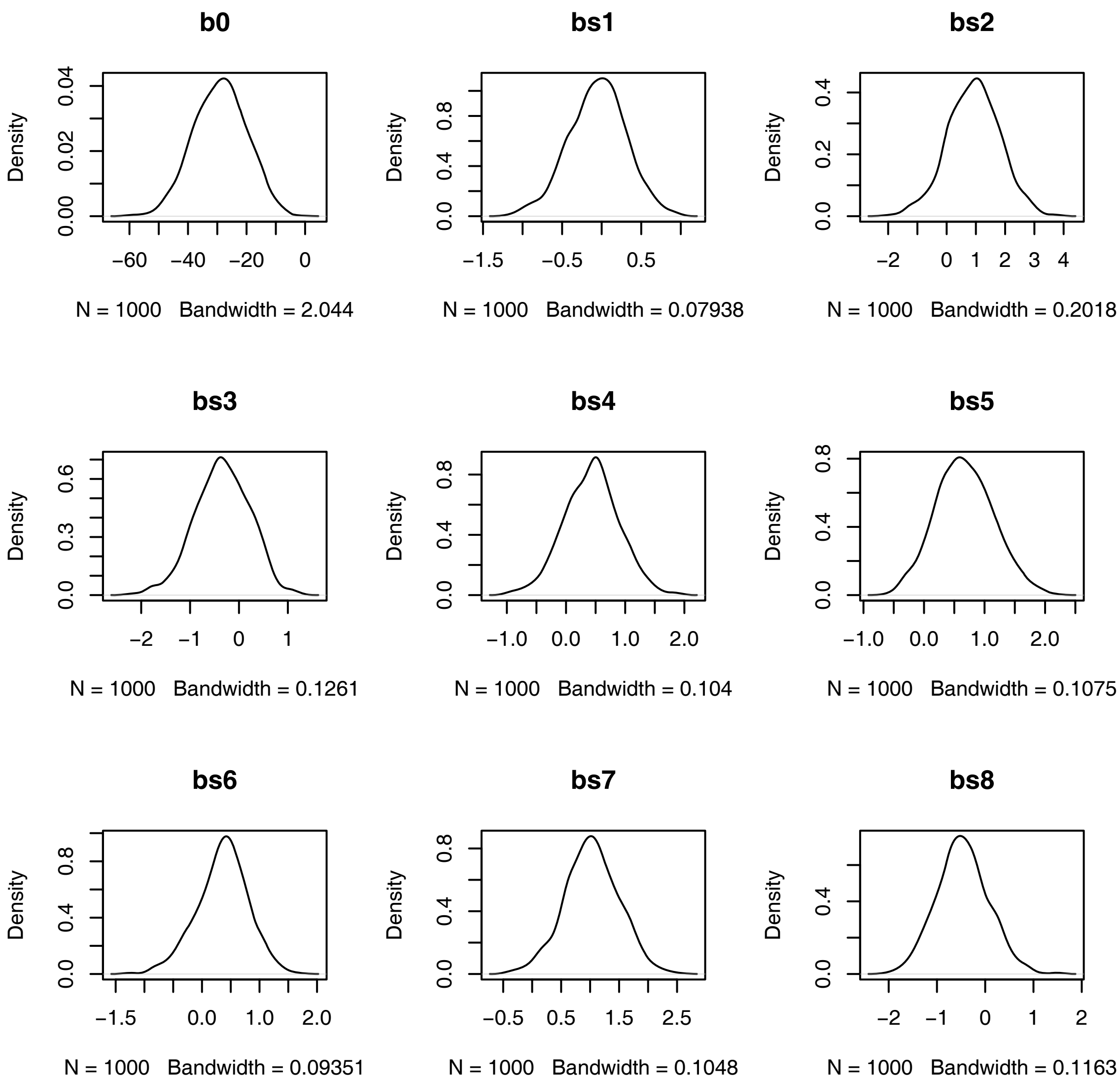

Figure 4.4: Plots of Posterior Densities-1

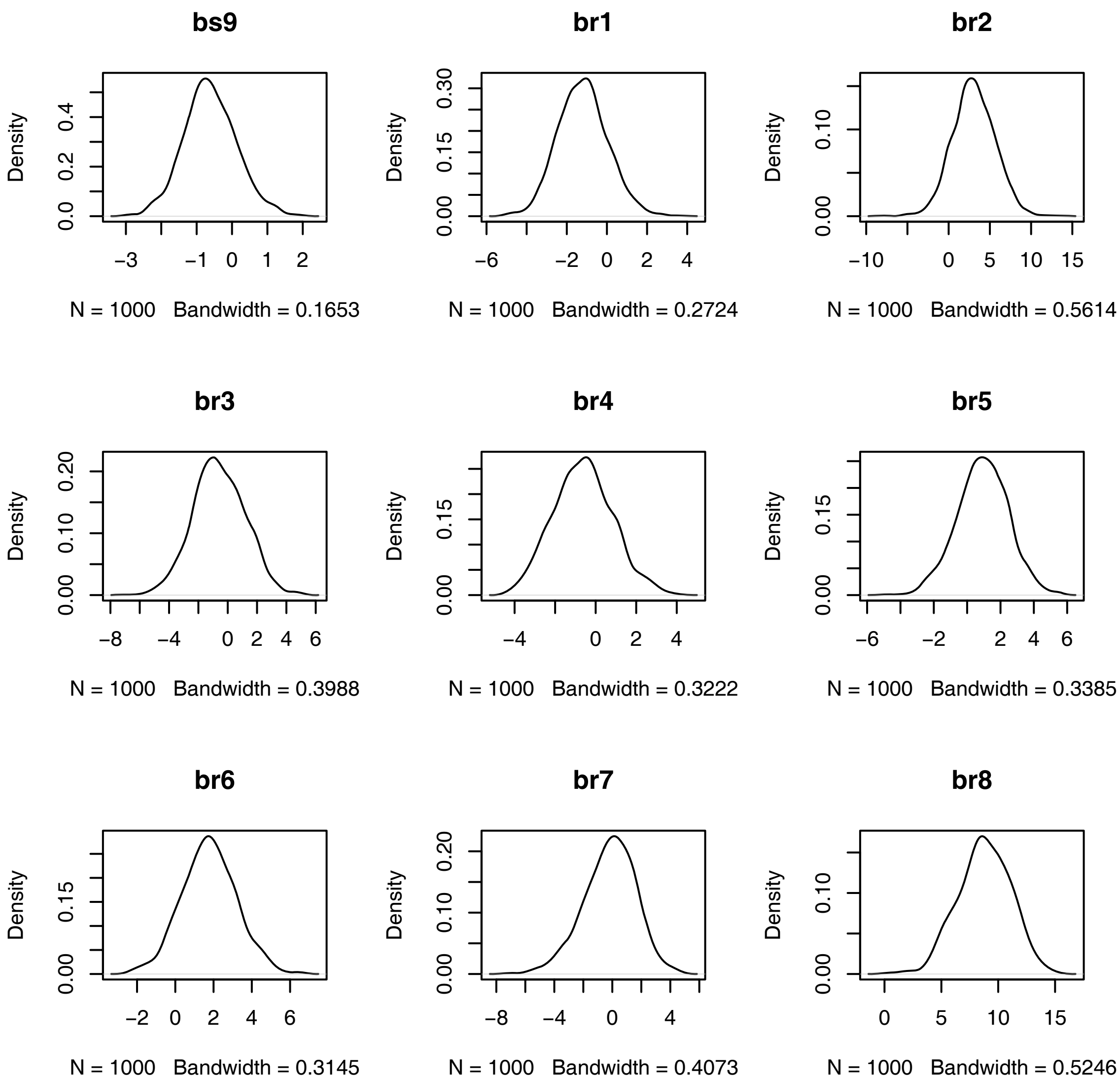

Figure 4.5: Plots of Posterior Densities-2

Chapter 5

Implications and Future Directions

This thesis focuses on investigating bidder characteristics using bilinear mixed models. Bilinear mixed model for network data has opened doors to many research opportunities in social sciences. We attempt to demonstrate one of its applications to investigate a complex behavioral process, i.e. overbidding in online auctions. Bidding data is converted to a bidder network framework to consider the second and third order dependencies. We consider the bidder level covariates to determine the characteristics of overbidders, which we term here as *Reactors*.

From the managerial perspective, this study provides interesting and useful insights for the auction house managers to support their revenue generating efforts. Our work illustrates the characteristics of *Reactors* in auctions. This information will be helpful to identify the registered bidders and create customized promotions to attract them to future auctions. Further, identifying bidders and tracking their bidding process will essentially allow the auction house managers to optimize the ordering of the lots and attain higher revenue. Although pre-auction estimates represent the expert's valuation of the item, an auction house may use it strategically to create a competitive bidding environment.

As with most research work, this study also has certain limitations that need to be acknowledged. First, it is difficult to determine the overbidders if they have not used a proxy bidding system. We used a very conservative measure of identifying such bidders by only considering those who have placed a higher normal bid in the auction after placing a proxy bid earlier. This measure is similar to one considered by Ku and his colleagues [11] in their study on auction fever. Further, there are certain types of bidders who use proxy bidding as a part of their bidding strategy and use it excessively (referred as *Agent Bidders* by Bapna et. al. [3]). Therefore, more advanced influence measures need to be considered in future studies for a better understanding of overbidders and *Influencers*.

This study considers a bidder's value change during an auction as a strategic element in designing auctions that are more effective. Therefore, sophisticated models may be developed in the future to predict which lots will attract overbidders and which will not. This may help in further optimizing lot orders for auctions. Our study contributes by investigating competition among bidders from the bidder level. Further studies on this topic are essential to understand bidding dynamics in auctions. Although it was not possible in this study, one may effectively identify overbidding or value change of bidders by taking a survey of bidders prior to the auction and record their pre-auction valuation for the items on which they intend to bid. This will essentially indicate whether a particular bidder has changed his valuation or not. Finally, further similar studies in auctions of other products like real estate will validate our findings on *Influencers* and *Reactors*.

We hope that this study will encourage other researchers to investigate the process of a bidder's value change more closely using other statistical techniques. We also hope that this will promote future applications of bilinear effects models and other advanced techniques in other areas of social sciences.

Bibliography

[1] Ariely, Dan and Simonson, Itamar (2003). Buying, Bidding, Playing, or Competing? Value Assessment and Decision Dynamics in Online Auctions. *Journal of Consumer Psychology* **13(1&2)** 113-123.

[2] Ashenfelter, O. and Graddy, K. (2002). Art Auctions: A Survey of Empirical Studies, National Bureau of Economic Research, *Working Paper*.

[3] Bapna, Ravi, Goes, Paulo, Gupta, Alok and Simonson, Itamar (2004). User Heterogeneity and its Impact on Electronic Auction Market Design: An Empirical Exploration. *MIS Quarterly* **28(1)** 21-43.

[4] Dass, Mayukh and Reddy, Srinivas K.(2007). Bidder Behavior and Bidder Networks in Online Auctions. *Working Paper* Athens, GA: University of Georgia.

[5] Heyman, James E., Orhun, Yesim and Ariely, Dan (2004). Auction Fever: The Effect of Opponents and Quasi-Endowment on Product Valuations. *Journal of Interactive Marketing* **18(4)** 8-21.

[6] Hoff, Peter D. (2005). Bilinear Mixed Effects Models for Dyadic Data. *Journal of the American Statistical Association* **100(469)** 286-295.

[7] Hoff, Peter D. (2005). gbme - R-functions for statistical analysis of social network data. *http://www.stat.washington.edu/hoff/Code/GBME/*

[8] Hoff, Peter D. and Ward, Michael D. (2004) Modeling Dependencies in International Relations Networks. *Political Analysis* **12(2)** 160-175.

[9] Kagel, John H., Levin, Dan, Battalio, Raymond C. and Meyer, Donald J. (1989) First-Price Common Value Auctions: Bidder Behavior and The "Winner's Curse." *Economic Inquiry* **27(2)** 241.

[10] Kagel, John H. and Richard, Jean-Francois (2001) Super-Experienced Bidder in First-Price Common-Value Auctions: Rules of Thumb, Nash Equilibrium Bidding, And The Winner's Curse. *Review of Economics & Statistics* **83(3)** 408-419.

[11] Ku, Gillian, Malhotra, Deepak and Murnighan, J. Keith (2005) Towards a Competitive Arousal Model of Decision-Making: A Study of Auction Fever in Live and Internet Auctions. *Organizational Behavior & Human Decision Processes* **96(2)** 89-103.

[12] Laudon, Kenneth C. and Traver, Carol G. (2004) . *E-commerce: Business, Technology, Society*, 2nd edition. Addison-Wesley, Boston, MA.

[13] MacEachern, Steven N. and Berliner, L. Mark (1994) Subsampling The Gibbs Sampler. *American Statistician* **48(3)** 188-190.

[14] Mei, J. and Moses, M. (2002). Art as an Investment and the Underperformance of Masterpieces, *American Economic Review*, **92(5)**, 1656-1668.

[15] Milgrom, Paul R. and Weber, Robert J. (1982) A Theory Of Auctions And Competitive Bidding. *Econometrica* **50(5)** 1089-1122.

[16] Narayan, V. and Yang, S. (2006). Trust between Consumers in Online Communities: Modeling the Formation of Dyadic Relationships. *Working Paper* New York: New York University.

[17] Park, Young-Hoon and Bradlow, Eric T. (2005) An Integrated Model for Bidding Behavior in Internet Auctions: Whether, Who, When, and How Much. *Journal of Marketing Research* **42(4)** 470-482.

[18] Reddy, Srinivas K. and Dass, Mayukh (2006) Modeling Online Art Auction Dynamics Using Functional Data Analysis. *Statistical Science* **21(2)** 179-193.

[19] Smith, Charles W. (1989). *Auctions: The Social Construction of Value*, The Free press, New York.

[20] Warner, R., Kenny, David A. and Stoto, Michael A. (1979) A New Round Robin Analysis of Variance for Social Interaction Data. *Journal of Personality and Social Psychology* **37** 1742-1757.

[21] Wasserman, Stanley and Faust, Katherine (1994). *Social Network Analysis: Methods and Applications*, Cambridge University Press, New York.

[22] Wong, George Y. (1982) Round Robin Analysis of Variance via Maximum Likelihood. *Journal of the American Statistical Association* **77** 714-724.