RICE UNIVERSITY

Structural Analysis of the "Best Offer" Mechanism on eBay

by

Kerem Yener Toklu

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Robin C. Sickles, Chair
Reginald Henry Hargrove Chair of Economics
Rice University

Mahmoud El-Gamal
Professor of Economics, and Chair in Islamic Economics, Finance, and Management
Rice University

Amit I. Pazgal
Professor of Marketing
Jesse H. Jones Graduate School of Business
Rice University

HOUSTON, TEXAS
APRIL, 2014
ABSTRACT

Structural Analysis of the "Best Offer" Mechanism on eBay

Kerem Yener Toklu

In this study, I examine the "Best Offer" mechanism on eBay by structural analysis. The methods I use in this study are based on the literature on structural auction and "Name your own price" models. Thus, in the first chapter, I provide a survey of this literature.

In the second chapter, I conduct a structural analysis, and propose a two step method to estimate the valuations of buyers who use the "Best Offer" format. The method is computationally attractive and can be implemented both parametrically and nonparametrically. It also has the flexibility to capture potential correlation across buyers’ valuations and different selling strategies. I test the performance of the estimator in Monte Carlo experiments, and find that the estimator performs reasonably well.
In the third chapter, I apply the estimation method to the used car market on eBay. I estimate buyers’ valuations and the determinants of them. I then run counterfactual experiments to find the optimal selling strategies. I find that the sellers’ behavior in the market is consistent with revenue maximization.

In the fourth chapter, I propose an alternative approach to estimate the distribution of buyers’ valuations based on weak behavioral assumptions. The method does not require the buyers to follow any particular offer strategy; therefore, it captures multiple equilibria that are possible to arise in the environment. Utilizing this feature, I test whether the valuation of a buyer who offers once is on average different from that of a buyer who makes multiple offers. The results suggest that sellers can be better off by following a harder bargaining strategy against the buyers who make multiple offers.
Acknowledgements

I would like to dedicate this work to my beloved Mom (Bahar) and Dad (Suleyman). I owe them everything I have and everything I have done.

I would like to thank my adviser, Robin Sickles, for leading and supporting me from the very beginning to this day.

I also would like to thank my committee members Mahmoud El-Gamal and Amit Pazgal for their advice and discussions.

Last but not the least, I would like to thank Seda Bulbul Toklu for accompanying me on this challenging journey. We did a lot together and shared a lot. What we went through was an exceptional experience.
Contents

ABSTRACT ii

Acknowledgements iv

List of Tables vii

List of Figures viii

Introduction 1

Chapter 1. A Survey of the Structural Analysis of Selling Mechanisms 5

Chapter 2. A Structural Model for "Best Offer" Sales on eBay 23

2.1. Introduction 23

2.2. The "Best Offer" Mechanism on eBay 26

2.3. Theoretical Model 28

2.4. Structural Econometric Analysis 34

2.4.1. Nonparametric Analysis 35

2.4.2. Parametric Analysis 39

2.5. Monte Carlo Experiments 42

2.5.1. Nonparametric Monte Carlo 43

2.5.2. Parametric Monte Carlo 47

2.6. Conclusion 50
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Parametric Monte Carlo</td>
<td>49</td>
</tr>
<tr>
<td>3.1</td>
<td>Summary Statistics for BO Listings</td>
<td>54</td>
</tr>
<tr>
<td>3.2</td>
<td>First Offers</td>
<td>55</td>
</tr>
<tr>
<td>3.3</td>
<td>Single Offers</td>
<td>56</td>
</tr>
<tr>
<td>3.4</td>
<td>Estimates for the Threshold Distribution</td>
<td>60</td>
</tr>
<tr>
<td>3.5</td>
<td>Structural Estimation Results</td>
<td>62</td>
</tr>
<tr>
<td>4.1</td>
<td>Estimation Results</td>
<td>82</td>
</tr>
</tbody>
</table>
## List of Figures

2.1 Estimated Joint Density of Values 44
2.2 Actual Joint Density of Values 45
2.3 Contour Plot of the Estimated Density 45
2.4 Contour Plot of the Actual Density 46
2.5 Marginal Density of Values 46
2.6 Inverse Offer Function 47
3.1 Estimated Threshold Density 61
3.2 Histogram of Actual Offers 64
3.3 Histogram of Simulated Offers 65
3.4 Bargaining Strategy with $\mu'=3\mu$ 68
3.5 Bargaining Strategy with $\mu'=2\mu$ 69
3.6 Bargaining Strategy with $\mu' = \mu$ 69
3.7 Bargaining Strategy with $\mu'=0.5\mu$ 69
3.8 Bargaining Strategy with $\mu' = 0.4 \mu$

3.9 Bargaining Strategy with $\mu' = 0.35 \mu$

3.10 Bargaining Strategy with $\mu' = 0.3 \mu$

3.11 Bargaining Strategy with $\mu' = 0.25 \mu$

3.12 Expected Seller Revenue vs Bargaining

3.13 Expected Seller Revenue vs BIN-Bargaining (Reservation Value = $0.8 \times \text{KBB Value}$)

3.14 Expected Seller Revenue vs BIN-Bargaining (Reservation Value = $0.6 \times \text{KBB Value}$)

3.15 Expected Seller Revenue vs BIN-Bargaining (Reservation Value = \text{KBB Value})

4.1 pdf of WTP for a high value car (KBB = $27,344)

4.2 cdf of WTP for a high value car (KBB = $27,344)

4.3 pdf of WTP for an average value car (KBB = $16,613)

4.4 cdf of WTP for an average value car (KBB = $16,613)

4.5 pdf of WTP for a low value car (KBB = $4,110)

4.6 cdf of WTP for a low value car (KBB = $4,110)
Introduction

Online markets have been studied extensively in the literatures of economics, management, and computer science since late 90s. eBay is one of the largest of them where almost everything from a pair of socks to private jets is traded today. To maintain a smooth trading environment, eBay has adopted different selling mechanisms, and updates these continuously to meet the varying demands of different parties. Auctions and fixed price sales are modified to create other selling formats such as auctions with "Buy it now" prices where buyers can either submit a bid for the ongoing auction or purchase the item immediately by paying the fixed price. Researchers have analyzed auctions and fixed price sales in great details, and "Buy it now" auctions to a lesser extent. They are compared from different perspectives including efficiency, profit maximization, consumer and producer surpluses\textsuperscript{1}. Aside from these sale formats eBay launched another hybrid selling mechanism called "Best Offer" in 2005. Though it has been highly used on eBay, previous research has not paid much attention to it. In this study, the "Best Offer" selling mechanism is analyzed by using a structural model. In particular, I propose a two step method to estimate buyers' valuations. I test the estimator in Monte Carlo experiments, and apply the method to a recently collected dataset of used cars listed on eBay Motors. I then run counterfactual simulations to examine the optimal selling strategies for the "Best Offer" mechanism.

\textsuperscript{1}See Hasker and Sickles (2010) for a recent survey on eBay.
Moreover, I also propose an incomplete model to estimate the distribution of buyers’ valuations under weak behavioral assumptions.

As an alternative selling mechanism, "Best Offer" (BO hereafter) is given as an option to sellers who choose to sell the item at a fixed price. Once they have chosen this option, sellers are able to receive offers from buyers. Buyers can either purchase the item by paying the fixed price or make offer(s) to the seller. The number of offers buyers can make is limited and offers expire after 48 hours. Sellers can accept, reject or let the offers expire. Moreover, a seller can make a counteroffer to a buyer if the buyer’s offer is not expired and the buyer’s offer limit is not reached. In this way the BO mechanism provides an environment for bargaining, an attractive feature for both sellers and buyers. Furthermore, sellers can automate their accept-reject decisions by entering thresholds in the system that cannot be observed by buyers.

The first chapter of this study provides a survey of the literature. Since the modeling approaches and econometric techniques are borrowed from structural auction and "Name your own price" models, I survey the papers in this literature. A wide variety of modeling assumptions and econometric techniques are discussed in this chapter.

In the second chapter, I propose a structural model to study the BO environment. The main goal in using a structural analysis is that it enables one to estimate the unobserved valuations of buyers, which cannot be obtained in a reduced form analysis. To the best of my knowledge, this is the first paper that uses a structural model to analyze the BO mechanism. I propose a two step estimation method that can be implemented both parametrically or nonparametrically depending on the data. The estimation method is based on the first order optimality condition of buyers,
and does not require the computation of the nonlinear offer function, which makes the estimation computationally attractive. I also allow for possible dependencies across buyers’ valuations and different bargaining strategies for the seller other than uniform bargaining, which is commonly assumed in the related literature. Monte Carlo experiments shows that the estimator performs well both parametrically and nonparametrically.

In the third chapter, I apply the proposed estimation method to a dataset of used cars sold on eBay Motors. I estimate the buyers’ valuations and determinants of them. I find that the variables signalling the quality of the car such as its book value or warranty are significantly positive. In contrast with some of the previous papers, I find that the feedback scores of sellers are insignificant as a determinant of buyers’ valuations. Using the calibrated model, I then run counterfactual experiments to examine the optimal selling strategies. I find that sellers’ behavior in the market is consistent with revenue maximization.

In the second and third chapters, I only focus on the buyers who chose to make a single offer for any listing\(^2\). Because most of the buyers make a single offer and most of the accepted offers are first offers, I do not lose the generality of the analysis focusing on single offers. Also, eBay does not provide the information of counteroffers by sellers publicly. So, it is difficult to analyze the following offers in the absence of this information since buyers possibly take counteroffers into account when they make subsequent offers. Moreover, the single offer focus in this study makes the model tractable enough for structural analysis.

\(^2\)I will use the terms "listing" and "sale" interchangeably from now on.
In the fourth chapter, I pursue the goal of estimating the distribution of buyers’ valuations under weak behavioral assumptions. The main advantage of this approach is that it allows one to capture many different equilibria that could occur in the BO environment since I do not impose any other restriction on the buyers’ decision process. It also enables the researcher to capture both single and multiple offers in a BO listing, an advantage that would be very difficult to obtain in a complete model as discussed in Chapter 2. However, the cost of this flexibility is the inability to point identify the valuations, since one can only identify and estimate the distribution of valuations. Since the analysis allows for the use of multiple offers, I then test whether there is any difference between the valuations of buyers who offer once and those who make multiple offers. I find that on average the valuations of multiple offerers is higher than those of single offerers for average or high value cars.
CHAPTER 1

A Survey of the Structural Analysis of Selling Mechanisms

In this chapter, I provide a summary of the studies that conduct structural analysis to examine different selling mechanisms. In particular, I study the structural approaches used for auctions and "Name your own price" mechanisms, since I use similar techniques to analyze the "Best Offer" format in the following chapters. I first start with the structural auction models, and discuss both parametric and non-parametric approaches for identification and estimation. I also discuss different modeling assumptions including private vs common values and participation of buyers. I then continue with the structural approaches used for the "Name your own price" mechanism.

Auctioning have been used as a selling mechanism since the antique ages. Today, many different goods and services including antiques and oil extraction leases are sold through auctions. With the advent of the online markets like eBay, auctions have become even more prevalent. In parallel with their importance, researchers have shown a significant interest in the topic, and studied many different aspects of it. Starting with the seminal paper by Vickrey (1961), early papers tried to establish a theoretical basis for auctions. Milgrom and Weber (1982) provide a general model for common auction types, and analyze bidding behavior and winning price. Wilson (1967) studies the information asymmetry across bidders and how it affects the bidding behavior.
Levin and Smith (1994) analyze the entry behavior of bidders and extend the revenue equivalence theorem accordingly.

From the beginning of 90s, researchers started to use the available bidding data to test the theoretical results in the auction literature. An important question in the empirical analysis of auctions is how to find the valuations or willingness-to-pay of bidders and their distribution. This is because knowing the valuations and their distribution, one can conduct counterfactual analysis to simulate the auction environment and provide answers to policy issues including optimal selling strategies, consumer and producer surpluses etc. However, the valuations are not observed in the real data except for the specific cases such as a second price sealed bid auction where the dominant strategy for each bidder is to submit one’s valuation truthfully. In general, the relation between the willingness-to-pay of a buyer and her bid can be quite complicated, and one cannot infer a buyer’s valuation trivially by observing her bid. This has generated the need to use the structural models to link the unobserved valuations with the observed bids. Since the computations involved in a structural analysis are burdensome, with the availability of faster processing computers, structural approaches have become more convenient.

One of the first structural papers on auctions is Donald and Paarsch (1993). They study a common problem in parametric estimation of structural auction models. In a classical first price sealed bid auction setting, the bidding strategy of a buyer is a nonlinear monotonic function of her valuation. If the distribution of buyers’ valuations is assumed to have a parametric form, one can derive the distribution of observed bids as a function of the parameters of the distribution of valuations. One can then try to
construct a likelihood function based on the observed bids and estimate the unknown parameters of the value distribution. The problem with this approach is that because the support of the bid distribution turns out to be dependent on the parameters of the value distribution, which is to be estimated, the regularity conditions of the maximum likelihood estimation are not satisfied. Hence, the standard asymptotic properties of the estimator may not be attained. To solve this problem, they propose a piecewise pseudo-maximum likelihood estimator. They show the conditions for its consistency and asymptotic distribution, and also show through Monte Carlo experiments that it performs well in small and medium size samples.

Another parametric estimation approach is proposed by Laffont et al. (1995) which is a simulation based method following McFadden (1989) and Pakes and Pollard (1989). In particular, they use a simulated nonlinear least squares approach which is adjusted to obtain consistent parameter estimates. They show the asymptotic properties of their estimator, and apply their method for a market of agricultural products which are sold through descending auctions. Their method also accounts for the heterogeneity across auctions and the case that only the winning bids are observed.

Although parametric models provided a good way for structural analysis, researchers have later attempted to explore less restrictive approaches. Guerre et al. (2000) propose a nonparametric estimation method for a first price auction environment with independent private values. Their method is implemented in two steps, and does not require the computation of the Bayesian Nash equilibrium strategy. In the first step, they nonparametrically estimate the distribution and density of the
observed bids. They substitute these estimates in the first order condition of bidders and estimate the private value of each bidder. In the second step, they nonparametrically estimate the distribution of bidders’ valuations. They also show that their estimator is uniformly consistent and attains the best possible convergence rate for this problem. To do this, they use the minimax theory developed by Ibragimov and Has’minskii (1981). Moreover, the method is computationally attractive since it does not require the computation of the equilibrium bidding strategies nor any optimization. Also, because the estimation is entirely nonparametric, it allows one to test the validity of the theoretical auction model.

When nonparametric estimation is to be utilized, an important issue is the identification of the model. Athey and Haile (2002) study this problem for the most common auction formats. They analyze different specifications including private vs common values, correlation across bidders’ valuations, and ex-ante bidder asymmetry. For the simplest independent private values model, they show that the transaction price is sufficient for nonparametric identification. They also show that an affiliated private or common values model is not identified if one or more bids are unobserved. Nevertheless, additional information such as bidder identities and bidder specific variables can make identification possible even in the absence of one or more bids. Furthermore, they show that the assumptions of the standard models can be tested even if the model is not identified.

Li et al. (2000) study a first price auction environment with correlated values. In particular, they analyze the case that the private information of each buyer is a product of two independent random signals. One of them is common to all bidders
and the other is a bidder specific signal. They consider two specific models based on different utility specifications. First, they study a conditionally independent private value model where the utility of each buyer is her private information, which are independently and identically distributed (iid) conditional on the common signal. Second, they study a pure common value model where the common signal becomes the same but unknown utility for all bidders. Utilizing the techniques from the measurement error literature, they show that the distributions of both signals can be identified for the private value model. They, however, make more assumptions to identify the common value model. They then propose a two step nonparametric estimation method where they use kernel estimators in the first step and characteristics functions in the second step. They also show the consistency of their estimator, and apply it to OCS Wildcat auctions. They find that the bidder specific signal explains the variability in bids more than the common signal does.

In another paper by the same authors, Li et al. (2002), they study affiliation in values, which is a specific form of dependency, and propose a nonparametric method to estimate the joint distribution of values. In particular, they consider an affiliated private value model, and extend the methodology in Guerre et al. (2000) for affiliated values. Hence, their estimator possesses the computational advantages of Guerre et al. (2000). Moreover, they show the consistency of their estimator and demonstrate its performance in Monte Carlo experiments. They also show how incorrectly assuming independency in values could change the results and policy implications.
Nonparametric estimation techniques enable a researcher to obtain robust estimates free from parametric assumptions. However, conducting a structural analysis also requires to make theoretical assumptions for the strategic behavior of the bidders. In general, English auctions are modeled as a "button auction", where bidders hold down a button to show their participation, and release it when the price reaches their valuations. This model captures the dominant strategy of truthfully bidding one’s valuation. Nevertheless, this approximation may not account for other important features of the bidding. In most English auctions, bidders call out their bids and the price increases in jumps; therefore, a bidder’s last bid may not reflect her actual valuation, and the model can perform poorly in explaining the data. Haile and Tamer (2003) take this problem in English auctions and propose an incomplete model which is based on two weak assumptions: 1- Bidders do not bid more than their valuations. 2- Bidders do not let an opponent win at a price they could beat. The assumptions are not only intuitive but also weak enough to cover different features of the bidding. It captures jump bidding and the case that bidders bid lower than their valuations. The assumptions are also satisfied in the equilibrium of the "button auction" model. These advantages, however, come at a cost which is the unidentification of the model. The assumptions allow for multiple equilibria and the valuations of the bidders are not uniquely identified. Haile and Tamer (2003) instead propose a method to estimate the bounds of the value distribution, and make inference accordingly. In particular, they suggest that each bid provides a lower bound on the valuation of the corresponding bidder, and each winning bid, up to the minimum bid increment, provides an upper bound on the valuation of the highest losing bidder. They then use the results
from order statistics to construct bounds for the distribution of valuations. They estimate these bounds nonparametrically, and show that they are tight. Moreover, if the "button auction" model is correct, then the bounds coincide and identify the value distribution. On the other hand, Haile and Tamer (2003) also show that if there is a small deviation from the "button auction" model, even if the two assumptions are satisfied, the estimated value distribution may not stay within the bounds. They also show how to construct bounds for the optimal reservation price and implement the method with covariates. Finally, they apply their method to US Forest Service timber auctions and examine the reserve price policies.

Most papers in the literature of structural auction models have examined private values, whereas only a few attempted to analyze the common value environment. This is mostly because of the complexities that arise in the common value analysis. Under the pure common value assumption, the bidding strategies become highly complicated even in the symmetric case. Therefore, researchers generally consider the special cases where the bidding strategies can be simplified. Smiley (1979) provides examples of these cases, and uses them to study oil and gas leases in the Outer Continental Shelf (OCS) in Gulf of Mexico.

Bajari and Hortacsu (2003) is one of the few papers that conducts an analysis with common values. They study coin auctions on eBay, and examine the determinants of buyer and seller behavior. For this purpose, they consider a second price common value auction model with endogenously entering bidders. They follow Levin and Smith (1994) and assume that bidders will continue to enter the auction until their ex ante expected utilities are zero. Hence, each bidder will play a mixed strategy for her
entry decision. Upon entry, bidders draw iid signals conditional on the common value of the auctioned object, and submit their bids accordingly. Their model also captures the "sniping" (last minute bidding) behavior of bidders in the sense that bidders do not want to bid early in the auction to avoid revealing any information about their private signals. As for the estimation strategy, the authors use a Bayesian approach for multiple reasons two of the which are as follows. First, Bayesian methods are computationally more convenient since classical methods such as maximum likelihood require optimization which may not perform well when the number of parameters are high. Second, as Donald and Paarsch (1993) point out, the standard asymptotic behavior of the maximum likelihood estimator is not guaranteed since the regularity conditions are not satisfied.

Another important phenomenon when working with common values is the "winner's curse". When a bidder is announced as the winner of an auction, this implies that the winner has the highest signal; in other words, he is the most optimistic bidder. This new information can be used to have a new expectation for the unknown common value which can be lower than the initial expectation of the winner, if the winner does not take this issue into account ex ante. This negative update in the expectations is called the "winner's curse". Explicit in its definition, winner's curse does not occur in equilibrium since bidders are assumed to behave rationally in equilibrium. Moreover, as the number of bidders increases, the extent of the winner's curse becomes larger. Bajari and Hortacsu (2003) use their calibrated model to estimate the winner's curse correction bidders make when they submit their bids. To do this, they calculated the change in the bid function when there is an extra bidder
in the auction. They also find the optimal minimum bid levels for the seller and the entry cost for the bidders, which can be the time and effort spent for bidding.

Another paper that uses a common value model is Hendricks et al. (2003). They analyze OCS Wildcat auctions which are used to lease tracts for oil and gas extraction in previously unexplored areas. In other words, bidders only observe seismic information but do not have any information from exploratory drilling in the area. The authors test whether bidders behave rationally implied by the equilibrium. They conduct multiple tests for this purpose. One advantage of their dataset is that they have an ex post variable to measure the value of the tract. More specifically, they constructed this variable by subtracting the discounted costs and royalty payments from the discounted revenues. This ex post value measure together with the bid data allow them to run multiple tests. First, they test whether the difference between the ex post value and the sale price is positive; namely, whether bidders make positive profit. They then test whether bidders bid less than what they expect ex ante, and also estimate the winner’s curse. In order to test the consistency of the bidding behavior with the equilibrium, they use the transformed first order condition as in Guerre et al. (2000). They find that bidders are aware of the winner’s curse, and bid in line with the equilibrium.

An important issue in structural auction models is the availability of the information on the number of bidders. This is because most identification results rely on that information; therefore, assuming that the number of bidders is known provides convenience in the analysis. However, this assumption may not be realistic for every environment including online markets. When there is a binding reserve price, only
bidders with values higher than the reserve price participate in the auction. Moreover, in an ascending auction, even such bidders may not find the chance to bid if the standing price is higher than their valuations. Hence, the observed number of bidders may not be a good proxy for the actual number of bidders. Song (2004) proposes an identification and estimation strategy for this problem. She uses the fact that the distribution of a random variable can be identified from the information on two consecutive order statistics i.e. the second and third highest bids. She then uses a semi-nonparametric estimation method proposed by Gallant and Nychka (1987). She tests the performance of her estimator in Monte Carlo experiments and apply it to the university yearbook auctions on eBay.

All the papers mentioned so far take each auction in an isolated manner ignoring the bidders’ problems in simultaneous or forthcoming auctions. When bidders are companies as in procurement or OCS auctions, they generally bid for multiple projects; therefore, they need to account for their involvements in other projects when bidding on any project, so that they could allocate their resources optimally. Considering this aspect, Jofre-Bonet and Pesendorfer (2003) study a dynamic auction game where they propose a method to estimate bidders’ private information. Following the method of Guerre et al. (2000), they base their estimation strategy on the first order optimality condition of bidders. Hence, their method does not require the computation of the equilibrium strategies which makes it computationally convenient. They then apply their method for the highway procurement contracts in California to find the effect of the previously taken and uncompleted projects (backlogs). Such projects can increase the cost of the current project through capacity issues; however, they
can also act in the opposite direction and provide more experience to the company thereby making it operate more efficiently and less costly. The authors conduct the structural analysis to find the net effect of the backlogs. They find that capacity constraints outweigh the experience effect, and raises the cost of the current project.

Another paper that examines the intertemporal dynamics in auctions is Sailer (2006). The author analyzes the intertemporal effects in personal digital assistant auctions on eBay, and argues that if these are not taken into account, bidders’ valuations are not estimated correctly. She approximates the eBay environment as an infinite horizon sequential Vickrey auction. Bidders participate in an auction if the expected return from it is higher than the participation cost. In general, the optimal strategy for a bidder in a static Vickrey auction is to submit her valuation truthfully. However, when the game is played sequentially, the optimal strategy changes. A participating bidder shades her valuation by the amount of her continuation value. Bidders are allowed to have different bidding costs. Bidders with high bidding costs would not want to bid multiple times; therefore, behave aggressively to win the current auction. Hence, the differences in the bidding costs yields the price dispersion, an analogy the author borrows from the search literature. Regarding the estimation strategy, Sailer (2006) adopts a multi step estimation method following Guerre et al. (2000). She first estimates the bid distribution extending the identification result of Song (2004) for asymmetric bidders. She then estimates the distribution of valuations up to location, winning bid, and bidders’ costs, respectively.

Most papers in the structural auction literature assume risk neutrality for bidders, although there is also evidence for risk averse behavior (Bajari and Hortacsu,
This is mainly because of the complications that arise when the risk neutrality assumption is relaxed. In particular, Guerre et al. (2009) show that the general first price auction model with risk averse bidders is not identified. Campo et al. (2011) examine this issue under alternative assumptions. They first show that the first price auction model is not identified even if the utility function is restricted to a constant absolute risk aversion (CARA) or a constant relative risk aversion (CRRA) model. The identification can neither be obtained when the value distribution is parameterized and the utility is left nonparametric. Seeking for the identification conditions, Campo et al. (2011) show that the model becomes identified when there is a parametric quantile restriction on the value distribution and the utility function is parametric. They then characterize the optimal rate that any estimator of the risk aversion can achieve utilizing the minimax theory by Ibragimov and Has’minskii (1981). Specifically, they find that the optimal convergence rate is independent of the number of covariates; hence, even though the convergence is slower than the parametric rate, the curse of the dimensionality is avoided. They then propose a multi-step semiparametric estimation procedure including nonparametric methods as well as weighted nonlinear least squares. They also study the extensions of their model allowing for binding reserve prices, affiliated private valuations, and asymmetric value distributions. Similar to Guerre et al. (2000), their method is advantageous in terms of computational conveniences. They apply their method to US timber auctions, and reject the risk neutrality.

Another important point in the analysis of auctions is the participation behavior of bidders. Some potential bidders might not find an auction attractive enough to
submit a bid; hence, they may choose not to participate in it. This might be due to high participation costs such as time and money needed to be spent to gather information about the auction and to prepare a bid. Ignoring this participation aspect might cause a selection bias in the analysis. Theoretical papers (Levin and Smith, 1994; Samuelson, 1985) have put emphasis on this issue and provide useful insights. Empirical papers that conduct structural analysis; however, have examined this problem to a lesser extent. Complications involved in such an analysis is an important reason for that. Li and Zheng (2009) study this issue for the procurement auctions in Texas, where less than half of the potential bidders participate in the auctions. To analyze this environment, Li and Zheng (2009) consider three different bidding models with endogenous entry. First, following Levin and Smith (1994), they consider a model where bidders play mixed strategies for entry yet they do not observe the number of participating (actual) bidders, contrary to Levin and Smith (1994). In their second model, they purify the entry strategies assuming that bidders draw different signals before participation. In the third model they study, bidders are assumed to draw their private costs before participation, similar to Samuelson (1985). One interesting result they find is that high number of potential bidders may cause less aggressive bids, in contrast with the well known previous results. In the standard independent private value model without entry, competition results in more aggressive bids. When entry is introduced, it counteracts with the well known "competition" effect and yields an opposite result. To quantify these effects and make policy recommendations, Li and Zheng (2009) estimates the structural parameters of entry and bidding. Because the necessary computations for such a task
is quite involved, they use a Bayesian approach and exploit the recent advancements in Markov Chain Monte Carlo techniques. They also allow for unobserved heterogeneity in their analysis, and find that it is significant for the procurement environment they study. They also find that the "entry effect" outweighs the "competition effect"; hence, a high number of potential bidders might increase the procurement cost of the government.

An interesting question in the auction literature is which auction format to use for a given market. Athey et al. (2011) analyze this issue for timber auctions. They especially study the effect of the auction format on bidder participation and competition. They collect data from both open and sealed bid auctions which are used to sell timber in Idaho-Montana border and California. The availability of the data from two different auction formats enables them to analyze multiple issues. They find that sealed bid auctions attracts small firms more than open auctions do; whereas, participation of large firms is similar in both auctions. They then consider a model capturing different aspects of the empirical evidence. They assume that bidders’ participation is costly and endogenous. Moreover, they allow for asymmetry in bidders to differentiate between small and large firms, and account for the possible collusive behavior of large firms. They estimate the structural parameters of the model parametrically using data from the sealed bid auctions and following the method by Guerre et al. (2000). They then use the calibrated model to predict the bidding behavior in open auctions. They find that the competitive bidding assumption fits well with the auctions in California; however, it hardly explains the price difference
in open and sealed bid auctions in Northern Forests. Furthermore, they also find that the data is consistent with the cooperative behavior of large firms at a mild level.

Most papers in the empirical auction literature assume that the covariate set captures all the information available to bidders. Nevertheless, this may not be a realistic assumption in many cases. The car market on eBay Motors is one example for that. The listing pages are highly informative since sellers can upload a lot of information including text, pictures, and video. Hence, it would generally be very difficult to capture all that information with the available covariates. The remaining information that is not controlled by the covariates cause unobserved heterogeneity, which should be accounted for in the analysis to obtain correct estimates. Pursuing this goal, Krasnokutskaya (2011) proposes a structural estimation method for a first price sealed bid auction under unobserved heterogeneity. She uses the results from the measurement error literature to estimate the distributions of valuations and unobserved heterogeneity. She also shows the uniform consistency of her estimator and applies it for the procurement auctions in Michigan. She then compares her results with those obtained under independence and affiliation assumptions, and finds that alternative assumptions overestimates the markup over bidders’ cost. Moreover, she also finds that the procurement costs obtained under IPV and APV assumptions are higher than those obtained under the unobserved heterogeneity assumption.

Although auctioning and fixed price sales have been popular selling formats, there are also other highly used selling mechanisms. One of them is the "Name your own price" (NYOP hereafter) mechanism launched by Priceline, where buyers make offers to buy flight tickets. The difference between a first price auction and NYOP is
that, in the latter, buyers do not compete with each other but try to make offers acceptable by the seller. Because of this difference researchers tend to use decision theoretic models for NYOP markets. Hann and Terwiesch (2003) study a German NYOP online retailer and quantify the effect of frictional costs which are defined as the costs incurred by the buyers making an offer. The retailer they examine initially adopted a single offer mechanism, and then switched to the multiple offer format. In the new format, buyers can increase their offers systematically to better guess the threshold price of the retailer; however, since each offer is costly they would need to optimize their offers and the number of offers they make. Hann and Terwiesch (2003) estimate the frictional cost and find that it is significant that fully exploitation of the information by buyers is not realized. They also find a learning experience that makes experienced buyers have a lower frictional cost.

Terwiesch et al. (2005) also study a German NYOP retailer. They model the consumer bargaining process in a stochastic dynamic programming setting. They derive the optimal threshold price for the retailer and compare it with the price used by the retailer. They show that the optimal price they obtained could significantly increase the retailer profits. They also examine the conditions under which haggling can be advantageous relative to posted prices. They find that if there is enough heterogeneity in the valuations and haggling abilities of consumers, haggling can be more profitable than posted prices.

As there have been different versions of the NYOP mechanism used by different online retailers, a natural question is whether any of these versions is more beneficial than others. Fay (2004) takes this issue from an interesting perspective. He compares
the single bid mechanism with the one that allows multiple bids. He finds that both mechanisms yield the same expected profit. Furthermore, he touches upon an important caveat of the single bid mechanism, which is the enforceability of the single bid rule. Because some buyers might sneak into the system by fake IDs, they might indeed submit multiple bids and take advantage of their previous bidding experience. Fay (2004) examines how this problem affects the profitability of the mechanism. He finds that profits are lower in this case compared to those from the benchmark mechanism. He also suggests that if such behavior cannot be avoided, a multiple bidding mechanism can be adopted.

When sellers use different channels to sell their products, the interaction of these channels can be critical. Sales through one channel can cannibalize the other; therefore, sellers need to optimize across different channels. Wang et al. (2009) take this issue and study different relevant aspects. They explore the conditions under which selling through the NYOP channel is profitable, and the optimal prices for the posted price and NYOP channels. They also examine the situation when the seller owns the NYOP retailer. To conduct their analysis, they consider a two stage game theoretic model with two types of customers: leisure and business travelers. Leisure travelers plan their trips in advance; whereas, business travelers make their arrangements close to the date of their travel, and they have higher valuations compared to leisure travelers. Wang et al. (2009) find that NYOP channel is profitable only if the available capacity does not largely exceed the expected number of business travelers and their valuations. They also claim that setting a price for the NYOP channel before the realization of the business travel demand is better. They conclude that the main
reason for using the NYOP channel is the uncertainty in the expected business travel demand, not the expectation of excess capacity.
CHAPTER 2

A Structural Model for "Best Offer" Sales on eBay

2.1. Introduction

In this chapter, I introduce the BO mechanism on eBay and provide a structural model to analyze the buying and selling behavior. In particular, I propose a two stage method to estimate buyers’ valuations and the determinants of them. This is important not only because it reveals the factors affecting valuations, but also enables the researcher to conduct counterfactual simulations on a variety of different research questions such as optimal selling strategies, consumer surplus etc. The method can be implemented both parametrically and nonparametrically depending on the availability of the data. It is also computationally easy to implement due to its two stage structure. Moreover, the method does not require the computation of the nonlinear offer function which brings further convenience in the estimation. Having discussed the estimation details, I then test both the nonparametric and parametric estimator in Monte Carlo experiments. Results show that the estimator performs reasonably well both parametrically and nonparametrically.

While the BO mechanism on eBay has not received much attention yet, the literature has rather focused on another similar selling mechanism called "Name your own price" (NYOP hereafter). The NYOP mechanism had been used by the online

\[1\text{See, for example, Wang et al. (2009), Terwiesch et al. (2005), and Amaldoss and Jain (2008).}\]
travel website "Priceline" before eBay launched the BO format\textsuperscript{2}. Today different versions of the NYOP mechanism exist for different markets. Buyers can make single or multiple offers in some markets, and in general travel websites use the NYOP format to sell "opaque" goods whose details are not revealed to buyers before purchase. An example of this is the sale of flight tickets for a particular route and date but with an unknown carrier and flight time. The used car market studied in Chapter 3 is Whereas "regular" in the sense that buyers see the item properties before taking their actions. The paper also makes technical contributions to the literature on the structural NYOP models. Two classical assumptions in this literature are that seller’s threshold distribution is uniform and buyers’ values are independent. Assuming these, one can obtain analytical expressions convenient for the structural analysis. Two of the papers that rely on these assumptions are Hann and Terwiesch (2003) and Spann et al. (2004). They estimate the cost of making offers and buyer valuations for NYOP sales, respectively. In this study, I conduct a structural analysis relaxing these two assumptions since these are restrictive for the BO environment. In particular, I propose a two step estimation method allowing for possible dependency across buyer valuations and more general threshold distributions. The estimation method is based on the first order optimality condition of buyers and does not require solving for the offer function which does not always have a closed form representation. Although, the analysis is for the BO mechanism on eBay, the methods can directly be applied for the NYOP mechanisms in other markets.

\textsuperscript{2}www.priceline.com uses this mechanism for flight bookings, hotel reservations, car rentals and other travel arrangements.
The proposed estimation technique is inspired from the two different streams of the literature: Structural auction models and NYOP models. I use a decision theoretic model as is generally used in the literature on NYOP mechanisms; furthermore, I followed the two stage estimation method proposed by Guerre et al. (2000) for first price sealed bid auction models. Early papers in the literature on structural auction models are Donald and Paarsch (1993) and Laffont et al. (1995). Both papers propose parametric estimation methods. On the other hand, Guerre et al. (2000) use the first order condition to identify and estimate the distribution of bidders' values nonparametrically without requiring the computation of the nonlinear bid function. Bajari and Hortacsu (2003) studies bidding behavior in coin auctions on eBay using a common value auction model with endogenous entry. Aside from finding the optimal reserve price, they also examine the extent of the winner’s curse. Nekipelov (2007) studies entry deterrence and learning prevention effects on eBay using a continuous time model and thus capturing within auction dynamics. Yao and Mela (2008) model both seller and buyer behavior in a structural model, and test how different strategies of the auction house affect the number of auctions and the bidding behavior.

The next section introduces the BO mechanism on eBay. The theoretical model is introduced in Section 2.3. Section 2.4 describes the econometric analysis. Monte Carlo experiments and their results are given in Section 2.5. Section 2.6 concludes.
2.2. The "Best Offer" Mechanism on eBay

eBay launched the BO mechanism in 2005. "Best Offer" option is given to fixed price sellers and not available in all the markets of eBay. Sellers can list their item for 3, 5, 7, 10 or 21 days under the fixed price format, and if they choose the BO option they can receive offers from buyers. Furthermore, throughout the sale buyers always have the chance to purchase the item immediately by paying the BIN price, and can avoid haggling in this way. Sellers can manually accept or reject offers, or can enter private thresholds in the system so that an offer above (below) the seller’s threshold is automatically accepted (rejected). Sellers can respond buyers by making them counteroffers. A transaction is made when an offer or counteroffer is accepted.

In eBay Motors, which is the market for motor vehicles and parts on eBay, buyers can make up to ten offers. This limit includes both offers a buyer makes and counteroffers she receives. Each offer is valid for 48 hours, and expires after that. While the listing is active, offer amounts are not revealed. So, buyers can see who has made offer(s) previously, whether these offers are pending, expired or declined, the date and time of the offers but not the offer amounts. Sellers, whereas see all the information. Offer amounts are then displayed on eBay when the listing ends with or without transaction.

Aside from the BO mechanism, there are other sale formats on eBay: auctions, fixed price (BIN) sales, and auctions with BIN prices. Of these mechanisms, in the

---

3Sellers using "Classified Ads" on eBay can also add the BO option. Buyers interested in items under "Classified Ads" contact the seller privately, and the transaction takes place off eBay.
4In other markets on eBay the maximum number of offers is three.
5In BO listings buyer IDs are not fully displayed. Characters in IDs except the first and the last ones are hidden by asterisks. Buyers can still figure out the number of their competitors from their partially hidden IDs. Feedbacks scores of buyers are also displayed.
car market on eBay, auctions including the ones with the BIN option constitute the majority of the listings. BO listings come after auctions and are more than fixed price listings. For instance, at anytime during the data collection process from eBay Motors\textsuperscript{6}, around 35-40 thousand vehicles are listed under the cars and trucks category. On average, more than half of these listings are auctions, around 20\% of them are BO listings, and fixed price listings constitute the smallest portion with less than 10\%. Moreover, comparing the ratio of successful sales in their dataset of Toyota Camrys across selling formats, Chen et al. (2011) find auctions w/o BIN option come first with 32.93\%, and BO listings have 18.31\% success ratio following auctions. Whereas, fixed price listings end up with sale with 4.71\%. Hence, considering the large trade volume on eBay, it is crucial to understand the BO mechanism, yet it has not been studied enough in the literature.

Furthermore, it is worth noting some other features of the BO mechanism. In order a buyer to win the listed item, she should either pay the BIN price or make an acceptable offer to the seller. So a buyer wins the item regardless of her competitors’ actions. Whereas in auctions with met reserve prices buyers compete each other to win the auction at the lowest possible price. The difference allows researchers to approach each mechanism in different ways. Strategic interaction between bidders in auctions is taken into account in game-theoretic models. Whereas, the bilateral interaction between a buyer and the seller can be modeled by decision theoretic approaches as has been done in the literature on the NYOP mechanism. In particular, sellers are

\footnote{http://motors.shop.ebay.com/Cars-Trucks-/6001/i.html?rt=nc\&__d=1\&__dmpt=US_Cars_Trucks\&__fso=0\&__fset=\&__jgr=0\&__qkw=1\&__sc=1\&__sop=1}
assumed to set private threshold prices, and accept (reject) offers higher (lower) than their threshold prices. Moreover, one might find that the BO mechanism is similar to the first price sealed bid auction format since in both mechanisms buyers submit the prices they are ready to pay. Indeed, this resemblance could be used to adapt the approaches used for first price auctions to BO sales if buyers’ offers were binding until the end of the listing. A similar idea is applied for merger models by Ivaldi and Motis (2007). They model the negotiations between a target firm, which is announced to be sold, and a number of acquiring firms as a first price auction where acquirers are taken as bidders. This can be a reasonable approximation since the target firm can choose the firm to merge with after negotiating all the acquirers, as a seller in a first price auction collects sealed bids first and announces the winner as the highest bidder. For the BO mechanism, however, such an approximation is not suitable. In particular, offers expire after 48 hours, so sellers cannot wait until the end of the sale to collect all offers from buyers and make a deal with the highest offerer. Nevertheless, if all buyers come in the last 48 hours, the mechanism turns out to be a first price sealed bid auction with a BIN price option, an interesting feature of the BO mechanism.

2.3. Theoretical Model

In this study I only focus on buyers who make a single offer. This is mainly due to the data limitations which is explained in detail in section 3.2. Most of the previous papers using structural models for auctions take seller behavior exogenously since knowing the value distribution and the bid function is sufficient to find the
optimal reserve price for the seller\textsuperscript{7}. However if, for example, one studies the reasons for selecting different sale formats i.e. auctions vs fixed prices, then sellers need to be modeled endogenously\textsuperscript{8}. Because I am interested in estimating the distribution of buyer valuations and optimal selling strategies under the given mechanism, I also take sellers’ side exogenously in this paper.

Another important point is that in the used car market on eBay, listing pages are quite rich in terms of information. eBay has a standard template for sellers to enter the features of the cars such as year, trim, mileage etc. In addition to this, eBay also allows sellers to post text or visual content including pictures and videos. Some sellers use the standard template whereas others use special software to design the listing page. Thus, listing pages are highly informative for buyers with all the text and visual content. Although, in my dataset I observe most features of the cars, I do not have all the information posted on listing pages. For instance, I do not have the interior details of the cars such as seats, audio system, airbags and so on. Therefore, the information that is not captured from the listing page may cause unobserved heterogeneity for the researcher. Furthermore, buyers also have the chance to contact sellers during the sale, and get additional information not posted on the listing page which also add to the unobserved heterogeneity in the data. As a result, even after conditioning on the observable covariates, buyers’ valuations can still be correlated through the unobserved information in the data. Hence, I have to account for this by

\textsuperscript{7}An example of these papers is Bajari and Hortacsu (2003).
\textsuperscript{8}See Baumer (2010) for example.
relaxing the independency assumption commonly used in the empirical literature on the NYOP mechanism.

For any listed item, buyers have two main strategies: paying the BIN price and purchasing it instantaneously, or making offer(s) to seller. I first consider the case where the BIN price is sufficiently high so that buyers never choose to exercise it, and then add the BIN option to the model.

I assume there exist finite numbers of risk neutral sellers and buyers in the market. For any listing with $N$ buyers, buyers draw their valuations from a commonly known continuous joint distribution $F(v_1, ..., v_N)$ with a corresponding density $f(v_1, ..., v_N)$ over the support $[\underline{v}, \overline{v}]^N$. I use the joint distribution to allow for possible correlation across buyers’ valuations in a listing. For each listed item, sellers choose a publicly observable BIN price and a private threshold price above which they are ready to accept offers. Naturally, seller’s threshold price should be lower than or equal to the BIN price. I assume sellers draw their threshold prices from a commonly known continuous distribution $G(.)$ with a corresponding density $g(.)$ defined on a bounded support $[\underline{x}, \overline{x}]$. Sellers accept offers above the threshold price, and decline or do not respond to offers below the threshold price. By definition, the upper bound of the threshold price, $\overline{x}$, should equal the BIN price, $p$, of the item. For the nonparametric analysis, I assume goods are homogeneous and no covariates exist. The analysis is extended to capture listing and seller covariates in the parametric analysis. Next, I pose the problem of a representative buyer with a value $v$, assuming that the BIN price is sufficiently high.
The buyer having a value $v$ makes an offer $b$ that maximizes her expected utility as follows:

\begin{equation}
U_O(v, G) = \max_b E[(v - b)I(x < b)|v]
\end{equation}

where, $x$ is the unobserved threshold price and $I(.)$ is the indicator function that equals 1 for acceptable offers and 0 otherwise. Equation (2.1) can be rewritten as $U_O(v, G) = \max_b (v - b)G(b)$. Taking the first order condition (FOC hereafter) and rearranging it gives,

\begin{equation}
v = b + \frac{G(b)}{g(b)} = \zeta(b, G)
\end{equation}

Equation (2.2) looks very similar to equation (3) in Guerre et al. (2000) who use it to identify and estimate the distribution of values nonparametrically in a first price sealed bid auction. The main difference between the two equations is that $G(.)$ and $g(.)$ here pertain to the threshold price which is independent of $v$, whereas in Guerre et al. (2000) they correspond to the bid variable which is a function of $v$. In Section 2.4, I propose an estimation method following the two stage approach in Guerre et al. (2000). I first estimate the threshold distribution, $G(.)$, and the density, $g(.)$, and then calculate pseudo values\footnote{This is the terminology used by Guerre et al. (2000) for the estimated buyer valuations.} for buyers and estimate the distribution of values. In their analysis of a NYOP market, Spann et al. (2004) assume a uniform distribution for threshold prices and obtain analytical expressions convenient for estimation. I
relax this uniformity assumption commonly used in the empirical NYOP literature, and allow for more general forms for the threshold distribution in the model.

I now discuss the existence and uniqueness of the offer strategy in the next lemma. To do this, I also assume a sufficient condition that $G(.)$ is log-concave as Guerre et al. (2000) discuss for the bid distribution. This condition is satisfied by most of the commonly used distributions such as normal, uniform, exponential\(^ \text{10} \).

**Lemma 1.** Let $G(.)$ be a log-concave distribution. It then follows that there exists a unique offer strategy, $b^* = s(v, G)$, that solves the buyer’s problem in equation (2.1). Moreover, the offer strategy is a strictly increasing function of valuation, $v$, and if $G(.)$ is strictly log-concave, then $\frac{db^*}{dv} < 1$.

**Proof.** First note that $b^* > v$ is not possible since buyer gets negative payoff. So it should be that $b^* \in [0, v]$. Also because the objective function in (2.1) is continuous with respect to $b$, there exists a $b^*$ that maximizes buyer’s payoff. Uniqueness of $b^*$ can be seen from the FOC in (2.2). Suppose there exists another maximizer $b^{**} > b^*$. Then, since $G(.)$ is log-concave, $\frac{G(b)}{g(b)}$ is nondecreasing, and thus $\zeta(b^{**}, G) > \zeta(b^*, G) = v$, a violation of the FOC. Similarly, $b^{**} < b^*$ is also not possible. Hence $b^*$ is unique.

The second part of the lemma can also be seen from the FOC. If buyer’s value, $v$, increases by one, $\zeta(b, G)$ should also increase by one. Because $\frac{G(b)}{g(b)}$ is nondecreasing, $\zeta(b, G)$ is strictly increasing in $b$. So $b$ should also increase. Furthermore, if $G(.)$ is

\(^ {10} \text{A distribution is (strictly) log-concave if its log is a (strictly) concave function. Bagnoli and Bergstrom (2005) discuss the properties and applications of log-concave distributions and densities in detail.} \)
strictly log-concave, then \( b \) can only increase by less than one since the change in
\[
\frac{G(b)}{g(b)}
\]
is also positive. Thus \( b^* = s(v, G) \) has a slope less than one.

Note that buyers always shade their values, \( s(v, G) < v \), to get positive expected payoff similar to bidding in first price auctions. So far I have assumed that the BIN price is sufficiently high so that it is not optimal for buyers to exercise it. Next, I relax this and show that buyers follow cutoff strategies such that buyers with high values may find it optimal to pay the BIN price and lower valued buyers make offers according to equation (2.2). Before presenting the next result I first define some additional notation for convenience. Let \( U_B(v, G) = v - p \) be the payoff buyer with a value \( v \) gets paying the BIN price, \( p \). Also let \( \bar{p} \) be the BIN price that makes the highest valued buyer indifferent between paying the BIN price and making an offer i.e. \( U_O(\bar{v}, G) = U_B(\bar{v}, G) = \bar{v} - \bar{p} \). So for all BIN prices greater than \( \bar{p} \) no buyer chooses to exercise the BIN option.

**Lemma 2.** Given a BIN price \( p < \bar{p} \), there exists a unique cutoff value \( v^c \in [p, \bar{v}] \) such that buyers with values \( v > v^c \) always exercise the BIN price, whereas buyers with values \( v < v^c \) choose to make an offer according to equation (2.2).

**Proof.** Both \( U_O(v, G) \) and \( U_B(v, G) \) are continuous functions of \( v \in [p, \bar{v}] \), and \( U_O(p, G) > U_B(p, G) = 0 \). Also, \( U_B(v, G) \) is an increasing function of \( v \) with a slope of one for all \( v \in [p, \bar{v}] \). By Envelope Theorem, \( \frac{\partial U_O(v, G)}{\partial v} = G(b(v)) < 1 \) since \( G(.) \) is a cdf. Therefore, there exists a cutoff value \( v^c \in [p, \bar{v}] \) such that \( U_O(v^c, G) = \).

\(^{11}\)Alternatively, it is enough to have \( \bar{v} \) sufficiently greater than \( p \).
Moreover, because the slope of $U_O(v, G)$ can never exceed one, $U_O(v, G)$ and $U_B(v, G)$ cannot intersect each other for $v > v^c$. Hence, $v^c$ should be unique. \hfill $\square$

Moreover, though it is explicit, it should be noted that only buyers with values greater than $x$ submit a positive offer as only bidders with values higher than the reserve price submit a positive bid in auctions. This will be discussed more in the next section.

### 2.4. Structural Econometric Analysis

In this section, I perform the econometric analysis for the model given in the previous section. Structural analysis enables one to infer about unobserved variables in the data and to make counterfactual experiments. In particular, it is essential, for example, to see the direct effects of listing and seller variables on buyer valuations. Moreover, to be able to compare different selling strategies for sellers, one should know how buyers respond to these strategies. Since in general seller’s alternative strategies are not observed in the data, one cannot use reduced form analysis for this purpose. The calibrated model can be used to simulate the seller’s revenue from alternative strategies and examine how optimally sellers behave in the market.

A crucial issue in estimating any structural model is identification of the model. In general, identification of structural auction models is based on the monotonicity of the bid function. Athey and Haile (2007) discuss most of the approaches used to identify and estimate auction models nonparametrically. Guerre et al. (2000) use the first order condition together with the monotonicity of the bid function to identify the distribution of values nonparametrically. Moreover, Li et al. (2000) and
Krasnokutskaya (2011) utilizes Kotlarski’s Lemma for nonparametric identification. There are also other identification results borrowed from "Competing Risk" models since both auction and competing risk models require the use of order statistics\textsuperscript{12}. In the next subsection, I discuss nonparametric identification and estimation for the model.

### 2.4.1. Nonparametric Analysis

The nonparametric identification results provided here are based on the arguments in Guerre et al. (2000) and Li et al. (2002). In a model for first price auctions with private valued bidders, they use the first order condition and the monotonicity of the inverse bid function to identify values nonparametrically. Their estimation method also follows the steps in identification and does not require deriving the bid function. This is a very attractive feature of their approach since in general bid function is highly nonlinear and standard parametric estimation techniques become computationally cumbersome. Following their way, I also use the first order condition in equation (2.2) to identify the model nonparametrically. Discussing the identification, I focus on one BO sale keeping the covariates fixed. Hence, I will keep suppressing the subindices for listings. Moreover, I assume that the BIN price is high enough so that exercising it is not optimal for any buyer i.e. $p > p_i$. Before giving the formal result I introduce some further notation for convenience following Athey and Haile (2007).

Let $\Psi$ denote the set of all continuous joint distributions over the observable offer variables. Also, let $\mathbf{b} = (b_1, ..., b_N)$ be the vector of offers in the listing and

\textsuperscript{12}See Rao (1992) for a nice summary of those identification results.
\( \mathbf{v} = (v_1, ..., v_N) \) be the corresponding vector of valuations. Offers have the lower and upper bounds, \( \underline{b} = s(\mathbf{v}, G) \) and \( \bar{b} = s(\mathbf{v}, G) \), respectively. The mapping \( \overline{\zeta}(., G) : [\underline{b}, \bar{b}]^N \rightarrow [\underline{v}, \bar{v}]^N \) is a vector of inverse offer functions such that:

\[
\overline{\zeta}(\mathbf{b}, G) = (\zeta(b_1, G), ..., \zeta(b_N, G)),
\]
given a threshold distribution \( G \). The structural model can then be defined as a pair \((\mathbf{F}, \gamma)\) such that \( \mathbf{F} \) is a set of continuous joint distributions over the unobserved values and \( \gamma : \mathbf{F} \rightarrow \Psi \) is a mapping such that \( \gamma(F') = \{H \in \Psi : H(b) = F'(\overline{\zeta}(b, G)) \ \forall b \in [\underline{b}, \bar{b}]^N\} \) for \( F' \in \mathbf{F} \). It is implicitly assumed that the model contains the true structure \((F, \gamma)\) where \( F \in \mathbf{F} \). The model is then said to be identified if, given the threshold distribution, the observed offers uniquely specifies the true structure in \((\mathbf{F}, \gamma)\). The next definition formalizes this idea.

**Definition 1.** Given a threshold distribution \( G \), the structural model \((\mathbf{F}, \gamma)\) is identified if for all \((F, \widetilde{F}) \in \mathbf{F}^2\), \( \gamma(F) = \gamma(\widetilde{F}) \) implies \( F = \widetilde{F} \).

**Lemma 3.** Let \( H_b \in \Psi \) be the joint distribution of observed offers \( b_i \) for \( i = 1, ..., N \). Given a threshold distribution \( G \), and the BIN price \( p > p^* \), the structural model \((\mathbf{F}, \gamma)\) is identified: \( \gamma(F) = \gamma(\widetilde{F}) = H_b \) implies \( F = \widetilde{F} \) for every \( F, \widetilde{F} \in \mathbf{F} \).

The proof of the lemma is straightforward since by definition of \( \gamma(.) \) and \( \overline{\zeta}(\mathbf{b}, G) \), \( \gamma(F) = \gamma(\widetilde{F}) \) implies \( F(\overline{\zeta}(\mathbf{b}, G)) = \widetilde{F}(\overline{\zeta}(\mathbf{b}, G)) \) for all offers, \( b \in [\underline{b}, \bar{b}]^N \). Thus, \( F = \widetilde{F} \) for every \( F, \widetilde{F} \in \mathbf{F} \).

\[\text{I assume } G \text{ as given since it can be estimated by using the rejected and accepted offers in the data. I discuss the estimation of } G \text{ in details in the parametric analysis section. Although I will assume a parametric form for } G, \text{ I refer to the estimation of value distribution as nonparametric here since I do not impose any parametric form on } F. \text{ One can also call it semiparametric to account for the parametric estimation of } G.\]
The joint distribution of values, $F$, can be estimated nonparametrically in two steps. In the first step $G$ is estimated by using the rejected and accepted offers in the data. I will give the details of it in the next subsection. For now I assume that $G$ and $g$ are estimated in the first stage by $\hat{G}$ and $\hat{g}$, respectively.

Let there be $T_N$ homogeneous BO sales with $N$ buyers who followed the single offer strategy\(^{14}\), and $\Omega_N$ be the set of these listings. Also, let $b_{it}$ be the offer of buyer $i$ in listing $t$. Given threshold distributions and densities, $\hat{G}_t$ and $\hat{g}_t$, and offers $b_{it}$ for $i = 1, ..., N$, and $t = 1, ..., T_N$ one can first obtain the pseudo-values $\hat{v}_{it}$ for buyer $i$ in sale $t$ from equation (2.2) as follows,

$$\hat{v}_{it} = \zeta(b_{it}, G_t) = b_{it} + \frac{\hat{G}_t(b_{it})}{\hat{g}_t(b_{it})}$$

Then, the joint density of values can be estimated by the simple kernel estimator as follows,

$$\hat{f}(v_1, ..., v_N) = \frac{1}{T_N h_f^N} \sum_{t=1}^{T_N} K_f \left( \frac{v_1 - \hat{v}_{1t}}{h_f}, ..., \frac{v_N - \hat{v}_{Nt}}{h_f} \right)$$

where $h_f$ is the bandwidth and $K_f$ is a kernel with compact support.

Because I have assumed $p > p$, no buyer chooses the BIN option. This allows the estimation of the pseudo-value of each buyer. When this does not hold, some buyers can choose to exercise the BIN price. The pseudo-values of these buyers can only be said to be greater than the critical value $v^c$. The above nonparametric identification result and estimation method can then only be applied for the truncated value density

\(^{14}\)Note that $N$ is observed for listings that ended without sale. Whereas, the number of observed buyers is only a lower bound for $N$ when the car is sold.
\( \hat{f}(v | v < v^c) \). In order to estimate the unconditional density, one needs to know the truncation probability \( F(v^c) \). If there is a consistent estimator of this probability, \( \hat{F}(v^c) \), the unconditional density can be estimated by \( \hat{f}(v) = \hat{f}(v | v < v^c) \hat{F}(v^c) \).

Suppose now that \( \Omega_N \) also contains the listings sold by the BIN option, and \( T_N \) is the number of listings in \( \Omega_N \). If \( N \) was known for all listings in \( \Omega_N \) including the sold ones, then one straightforward estimator of \( F(v^c) \) could be:

\[
\hat{F}(v^c) = \Pr(v_1 < v^c, ..., v_N < v^c) = \frac{\sum_{t=1}^{T_N} I(t \in S) + I(t \in NS)}{T_N}
\]

where, \( NS = \{ t \in \Omega_N : \text{listing } t \text{ ended up without sale} \} \) and \( S = \{ t \in \Omega_N : \text{\( N \)th offer in listing } t \text{ is accepted}^{15} \} \). However, this may not be a feasible estimator for two reasons. First, for sold listings the number of observed buyers is only a lower bound for \( N \). Thus, to apply the above method one would need to have more information to proxy for \( N \). One possible way to solve this issue could be to use the listings where an accepted offer is made close to the end of the sale. If buyers arrive listings homogeneously through time, then the probability of the arrival of another buyer in those listings would be less than that of other sold listings. Hence, the number of observed buyers would be a better estimate for \( N \). A similar approach is employed by Song (2004) to estimate the third highest valuation by the third highest bid for yearbook auctions on eBay. The second complication is that the critical value, \( v^c \),

\(^{15}\)If a \( j^{th} \) offer is accepted for \( j < N \), then one cannot say anything about the valuations of the \( N - j \) buyers who did not offer. Listings with such offers should also be left out of \( \Omega_N \).
depends on the covariates\textsuperscript{16} including the BIN price. Thus, one needs to have a large sample that contains listings with very close covariates. Considering the huge data available on eBay, this issue can also be overcome.

On the other hand, throughout the entire discussion above, it is implicitly assumed that the lower bound of $G$ is zero. Otherwise, a positive lower bound would serve as a reservation price in auctions, and create another truncation from below which would bring further complication into the model.

\textbf{2.4.2. Parametric Analysis}

Nonparametric analysis may not be feasible because of the complications mentioned in the previous subsection. Furthermore, the listing pages are highly informative; therefore, one would need to use a number of covariates when analyzing an eBay sale, which would be another problem for nonparametric estimation. As the number of covariates increases, nonparametric estimates converge at a lower rate and perform poorly, which is a well known issue called the "curse of the dimensionality" (Li and Racine, 2007). Hence, parametric estimation may be more appropriate for the analysis. I follow the same two stage estimation strategy based on the first order condition, but this time imposing a parametric structure on the value distribution, $F(.)$\textsuperscript{17}. In the first step, threshold distribution and density, $G(.)$ and $g(.)$, are estimated. Then

\textsuperscript{16}Note that this issue is different from the curse of the dimensionality in nonparametric estimation. Having a sufficiently large sample one can handle the latter problem and estimate the distribution of values conditional on the covariates.

\textsuperscript{17}Semi-parametric methods such as a single-index model by Powell et al. (1989) or semi-nonparametric approaches by Gallant and Nychka (1987) can also be used to estimate the distribution of valuations.
by using the estimates, $\hat{G}(,)$ and $\hat{g}(,)$, *pseudo-values* are obtained from equation (2.2), and $F(.)$ is estimated in the second step.

Let $Z = \{Z_t\}_{t=1}^T$ denote the set of observable covariates for all listings. The distributions of valuations and sellers’ thresholds may depend on these covariates. Let $F(.|Z_t)$ and $G(.|Z_t)$ be these distributions, respectively, conditional on the covariates. In order to operationalize Lemmas 1 to 3, I assume $G(.|Z_t)$ is a log-concave distribution with a parameter vector $\alpha_g$. Note that sellers’ thresholds are private information and unobserved to the researcher. To be able to estimate the parameter vector $\alpha_g$, one would have to use other observable information. I use the the offer data for this purpose. In particular, I use the information whether an offer is accepted or rejected. Because sellers reject offers below their thresholds and accept the ones above their thresholds, one can infer whether an offer is higher or lower than the seller’s threshold seeing the status of the offer. I use this set of information to construct a likelihood function, and estimate the parameter vector $\alpha_g$ by maximum likelihood principle. In the second step, having estimated the *pseudo-values* of buyers, I use the maximum likelihood method again to estimate the parameters of the value distribution.

Let $T$ be the number of listings in the sample and $\Omega$ be the set of the listings. The offer data can be partitioned into three mutually exclusive sets for convenience: $A_1 = \{t \in \Omega : \text{only rejected offers exist in sale } t\}$, $A_2 = \{t \in \Omega : \text{both rejected offers and an accepted offer exist in sale } t\}$, and $A_3 = \{t \in \Omega : \text{only an accepted offer exists in sale } t\}$. Because the threshold is assumed to be fixed throughout the sale, only

\[\text{Note that listings that ended through the BIN option without any offer are not useful for this purpose.}\]
the largest rejected offer is needed when the number of rejected offers is more than one. By definition, only one accepted offer can exist in any sale. Since thresholds are unobserved, the likelihood function consists of the product of the probabilities for observing intervals. Let $b_a^t$ and $b_r^t$ be the accepted offer and the maximum rejected offer in sale $t$ (if they exist), respectively. Representing the entire offer data by $b = \{b_a^t, b_r^t\}_{t=1}^T$, and the covariate set by $Z = \{Z_t\}_{t=1}^T$, the log-likelihood function can then be constructed as follows:

$$\log L_1(\alpha_g; b, Z) = \sum_{t=1}^T \{ I(t \in A_1) \log (G(p_t|Z_t) - G(b_r^t|Z_t)) + I(t \in A_2) \log (G(b_a^t|Z_t) - G(b_r^t|Z_t)) + I(t \in A_3) \log (G(b_a^t|Z_t) - G(x_t|Z_t)) \}$$

The estimates of the parameter vector $\alpha_g$ can then be obtained as,

$$\{\hat{\alpha}_g\} = \arg \max_{\alpha_g} \log L_1(\alpha_g; b, Z)$$

Having estimated the threshold distributions and densities, $\widehat{G}_t(.)$ and $\widehat{g}_t(.)$, I can now plug them back into equation (2.2) together with single offers to get pseudo-values of buyers. Let $v_{it}$ denote the pseudo-value of buyer $i$ who submitted a single offer in sale $t$. Also let $N_t$ be the set of these buyers in sale $t$. For notational convenience, I use $v_t = \{v_{it}\}_{i=1}^{N_t}$ to denote the set of valuations of the buyers in listing $t$. The distribution of valuations, $F(.|Z_t)$, depend on the observable covariates, $Z_t$, through the parameter vector $\alpha_f$ to be estimated. Since buyers can also choose to exercise
the BIN price, one should take it into account in the estimation to prevent a possible selection bias. For those buyers, one cannot point identify their valuations, but only infer that their valuations are higher than the critical value, $v_t^c$. I then construct the following likelihood function accounting for these points:

$$\log L_2(\alpha_f; Z, v) = \sum_{t=1}^{T} \{I(t \in NB) \log f(v_t|Z_t) + I(t \in B) \log (1 - F(v_t^c|Z_t))\}$$

where, $B$ is the set of listings where BIN price is exercised without any buyers making an offer\(^\text{19}\). $NB$ contains the listings where it is not exercised. The parameter vector $\alpha_f$ can then be estimated by:

$$\{\hat{\alpha}_f\} = \arg \max_{\alpha_f} \log L_2(\alpha_f; Z, v)$$

2.5. Monte Carlo Experiments

In this section, I test the estimators proposed in section 2.4.1 and 2.4.2. To do this, I conduct two Monte Carlo experiments simulating the "Best Offer" environment. First, I test the nonparametric estimator for a BO environment with homogenous listings. Second, I test the parametric estimator with heterogenous listings with covariates.

\(^\text{19}\)I only take BIN sold listings without any offers because eBay does not reveal the identity of the buyer who exercises the BIN price. So, one cannot know whether the buyer has made an offer before choosing the BIN price. Hence, including such listings might induce specification error in the analysis.
2.5.1. Nonparametric Monte Carlo

In this first experiment, I test the performance of the nonparametric estimator proposed in section 2.4.1. I assume that listings are homogeneous with no covariates. In each replication, I generate 500 BO listings with 2 buyers. In general, online markets are very rich in terms of data; therefore, a sample size of 500 is representative. The assumption of 2 buyers also comes from the data\textsuperscript{20}. In each listing, I assume that buyers’ valuations come from a bivariate normal distribution with the following mean vector and covariance matrix:

\[
\mu = \begin{pmatrix} 15 \\ 15 \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} 20 & 10 \\ 10 & 20 \end{pmatrix}
\]

Note that covariance terms are nonzero which makes the valuations of buyers correlated. I also assume that BIN price is 35, and sellers draw their private thresholds from a normal distribution with a mean of 25 and a standard deviation of 8, truncated from below at 0 and above at the BIN price. I first calculate the critical value above which a buyer would exercise the BIN price. Then, I draw the buyers’ valuations in each listing. If the first buyer’s valuation is higher than the BIN price, she exercises the BIN option; otherwise makes an offer according to the FOC. If the offer is accepted or the BIN option is chosen, then the listing ends without the arrival of the second buyer. If her offer is rejected, then the second buyer arrives, and either makes an offer or exercises the BIN option depending on her valuation.

\textsuperscript{20}In the data used in Chapter 3, the average number of buyers is 2.25 for listings where a single offer exists.
Having generated the data, I first estimate the parameters of the sellers’ threshold distribution as described in 2.4.2. In the second step, I estimate the pseudo-value of each buyer using the FOC, and then estimate the joint and marginal densities of valuations as in 2.4.1. To operationalize this, I follow Guerre et al. (2000), and use a triweight kernel \((35/32)(1 - u^2)^3 I(|u| \leq 1)\) and the bandwidths \(h_j = 1.06\hat{\sigma}_v(N_j)^{-1/5}\) and \(h_m = 1.06\hat{\sigma}_v(N_m)^{-1/5}\) to estimate the joint and marginal densities of valuations, respectively. \(\hat{\sigma}_v\) is the standard deviation of pseudo-values in the sample, and \(N_j\) and \(N_m\) are the number of listings and number of offers, respectively. The coefficient 1.06 comes from the so-called "rule of thumb" (See Silverman, 1986). I replicate this procedure for 1000 samples, and results are given in Figures 2.1 to 2.6.

Figure 2.1. Estimated Joint Density of Values

Estimated joint density of valuations and the actual density are given in Figures 2.1 and 2.2. The nonparametric estimator performs well and follows the actual density closely over its entire support. It also captures the mean, variance, and the correlation
across valuations well, as depicted in Figures 2.3 and 2.4. For practical purposes, it is also important to know about the marginal distribution of valuations. Figure 2.5, provides the nonparametric estimate of the marginal density obtained from the
pooled data of valuations. The actual density and the 95% confidence band are also given in the same Figure. The estimate of the marginal density also performs well over the entire support including the tails. The confidence band also contains
the actual distribution over the entire support. In Figure 2.6, the estimated inverse offer function is plotted and compared with the actual function together with a 95% confidence band. The inverse offer function almost coincides with the actual one for most of the values in its support. It deviates from the actual one for offers close to the BIN price. Confidence band also gets larger in the same interval. Since the inverse offer function is estimated by using the first stage estimate of the threshold distribution, the slight deviation in the inverse offer function stems from the error in the first stage estimation.

2.5.2. Parametric Monte Carlo

When the data is not large enough to make the nonparametric estimation feasible, one may have to conduct the analysis parametrically. In the second experiment, I test the performance of the parametric estimator in section 2.4.2, when there are
both continuous and discrete covariates. To make the simulation more realistic, I use covariates and parameter values consistent with the actual data of used Ford Mustangs. I assume that there are 2 buyers in each sale, and buyers in sale \( t \) draw their valuations from a multivariate normal distribution with a mean vector and covariance matrix specified as follows:

\[
\mathbf{v}_t \sim N(10^4 \mathbf{\mu}_t, 10^8 \Sigma_t), \quad \text{where} \quad \mathbf{\mu}_t = \alpha_1 + \alpha_2 KBB_t/10^4 + \alpha_3 PF_t/100 + \alpha_4 NF_t \quad \text{and} \\
\Sigma_t = \begin{bmatrix} \sigma_{11}^t & \sigma_{12}^t \\ \sigma_{21}^t & \sigma_{22}^t \end{bmatrix}, \quad \sigma_{11}^t = \sigma_{22}^t = \alpha_5 KBB_t/10^4 \quad \text{and} \quad \sigma_{12}^t = \sigma_{21}^t = \alpha_6 KBB_t/10^4.
\]

\( KBB_t, PF_t, \) and \( NF_t \) are the book value of the car, and numbers of positive and negative feedbacks the seller has received in the past 12 months, respectively. Note that cross diagonal terms in \( \Sigma_t \) represent the covariance across buyers’ valuations. I also assume that the seller in listing \( t \) draws her private threshold from a truncated normal distribution with a mean and variance linear in the book value of the car: \( X_t \sim \text{Trunc.Norm}(10^3 \mu_t, 10^3 \sigma_t) \) over the support \([0, p_t]\), where \( X_t \) denotes the threshold and \( p_t \) is the "Buy it now" price of the car, and \( \mu_t = \beta_1 + \beta_2 KBB_t/10^3 \) and \( \sigma_t = \beta_3 + \beta_4 KBB_t/10^3 \) are the mean and standard deviation of the threshold, respectively. I use the following parameter values for the simulation: \( \alpha = \{0.2543, 0.9537, 0.0367, -0.0463, 0.4676, 0.3391\} \) and \( \beta = \{1, 0.7, 0.2, 0.1\} \). Moreover, to generate the covariates, I assume that they come from a multivariate normal distribution, and I estimate the parameters of this distribution by maximum likelihood using the actual data. To wit, let \( Z_t = \{KBB_t/10^4, PF_t/100, NF_t\} \) denote the covariate vector for listing \( t \), and \( Z_t \sim N(\gamma, \Xi) \). \( \gamma \) and \( \Xi \) are

\(^{21}\)See subsection 3.3.2 for details of the data.
estimated as $\hat{\gamma} = \{1.6642, 0.5926, 0.3125\}$ and $\hat{\Sigma} = \begin{bmatrix} 0.7350 & -0.2725 & -0.1294 \\ -0.2725 & 6.5081 & 3.6933 \\ -0.1294 & 3.6933 & 2.8565 \end{bmatrix}$.

The reason for using a joint distribution instead of drawing from their marginals is to capture the potential correlation across the covariates. I also generated the price for each listing in a similar manner assuming that the price is normally distributed around the book value of the car: $p_t = KBB_t + \epsilon_t$, where $\epsilon_t \sim N(5000, 8200)$ is a normal random variable with the parameters in parenthesis showing the mean and the standard deviation. For each listing, I simulate the BIN price and the private threshold of the seller as described. I then draw the valuations of the two potential buyers. The first buyer makes an offer if her valuation is lower than the critical value, otherwise exercises the BIN price. If there is no sale, the second buyer arrives and behaves accordingly. To test the small sample properties of the estimator, I run the Monte Carlo experiments for sample sizes of 100, 200, and 300. The results are given in Table 2.1.

<table>
<thead>
<tr>
<th>Parameters $\alpha$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
<th>$n = 300$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_1$ = 0.2543</td>
<td>0.5293</td>
<td>0.4351</td>
<td>0.3964</td>
</tr>
<tr>
<td></td>
<td>(0.7620)</td>
<td>(0.1768)</td>
<td>(0.1375)</td>
</tr>
<tr>
<td>$\alpha_2$ = 0.9537</td>
<td>0.9997</td>
<td>0.9825</td>
<td>0.9867</td>
</tr>
<tr>
<td></td>
<td>(1.0339)</td>
<td>(0.0314)</td>
<td>(0.0198)</td>
</tr>
<tr>
<td>$\alpha_3$ = 0.0367</td>
<td>0.0271</td>
<td>0.0451</td>
<td>0.0387</td>
</tr>
<tr>
<td></td>
<td>(0.1701)</td>
<td>(0.0109)</td>
<td>(0.0053)</td>
</tr>
<tr>
<td>$\alpha_4$ = -0.0463</td>
<td>-0.0538</td>
<td>-0.0527</td>
<td>-0.0512</td>
</tr>
<tr>
<td></td>
<td>(0.2874)</td>
<td>(0.0239)</td>
<td>(0.0087)</td>
</tr>
<tr>
<td>$\alpha_5$ = 0.4676</td>
<td>1.6216</td>
<td>0.7709</td>
<td>0.6681</td>
</tr>
<tr>
<td></td>
<td>(97.4502)</td>
<td>(1.6899)</td>
<td>(0.7082)</td>
</tr>
<tr>
<td>$\alpha_6$ = 0.3391</td>
<td>1.3112</td>
<td>0.6281</td>
<td>0.5366</td>
</tr>
<tr>
<td></td>
<td>(63.4756)</td>
<td>(1.5464)</td>
<td>(0.6631)</td>
</tr>
</tbody>
</table>
For a sample size of 100, although parameter estimates for the mean parameter are close to their actual values, mean square errors (MSE) are quite large. As we increase the sample size, both parameter estimates get closer to their actual values and MSE decrease. An exception of this is the parameter of the book value of the car denoted by $\alpha_2$, which does not monotonically converge to its true value for the given sample sizes. On the other hand, the parameters pertaining to the covariance matrix, $\alpha_5$ and $\alpha_6$, converge relatively slowly. This shows that at least a medium size sample would be necessary to estimate the potential correlations across valuations. Moreover, sample sizes larger than 200 seem to provide plausible results for other parameters. Given the availability of the large data in online market places, the estimator would perform well in practice.

2.5.3. Conclusion

This chapter provides a structural estimation method for BO sales on eBay which allows researchers to estimate the unobserved valuations of buyers and conduct counterfactual analysis for policy analysis. The method can be implemented both parametrically or nonparametrically. It is based on the first order optimality condition of buyers and does not require the computation of the nonlinear offer function. This feature of the method makes it computationally advantageous compared to likelihood based approaches. The performance of the estimator is tested in Monte Carlo experiments. The estimator performs reasonably well both parametrically and non-parametrically.
CHAPTER 3

An Empirical Analysis of the Used Car Market on eBay

3.1. Introduction

In this chapter, I apply the estimation method, which is proposed in the previous chapter, for the used car market on eBay. In particular, I analyze the market of the used Ford Mustangs sold through the BO format. There are two reasons for my choice of this particular market. First, Mustangs can be considered as a premium segment; thus, substitution effects across sales would be minimal\(^1\). This would enable me to analyze each BO listing separately, thereby making structural methodology in the previous chapter convenient. The second reason is the availability of the data. In general, Mustangs are sold the most in the used car market on eBay; therefore, one would have a relatively larger sample size.

Although there are both theoretical and empirical papers on other sale formats, BO sales on eBay has not been studied as much. To the best of my knowledge, this is the first paper that studies the BO selling mechanism with a structural model. Also, together with Chen et al. (2011), the paper brings the only empirical evidence from eBay. Chen et al. (2011) analyze buyers’ behavior in a sequential game setup and provide empirical evidence from reduced form analysis consistent with their model.

\(^1\)On the contrary, substitution effects would be much higher for a lower segment such as Toyota Corolla, since potential buyers in this segment would be more price sensitive and more interested in the functionality of the car rather than its image or brand value.
In this chapter, however, I conduct structural analysis to identify the factors affecting buyer valuations and find out whether sellers choose their selling strategies optimally. For each BO listing, sellers first choose a "Buy it now" (BIN) price and then private thresholds or a bargaining strategy for possible offers. Using the calibrated model, I simulate the seller’s revenue for alternative strategies, and find that sellers’ behavior in the market is consistent with revenue maximization.

I collected the data using a spider program coded in Java which visits both the active and completed listings to collect different sorts of information. For each listing, I observe car attributes as well as the seller variables such as feedback scores. I also collected the book values of the cars from the Kelley Blue Book website. Because there are a number of covariates in each listing, I conduct the analysis parametrically.

I first estimate the seller’s threshold distribution using the offer data, in particular the information whether an offer is accepted or not. In the second stage, I estimate the valuations of buyers and their joint distribution. I find that the book value of the car, auto check score, and the warranty of the car have a significantly positive effect on the valuations. Contrary to some other studies, I find that feedback scores are not significant. This might be because of the generality of the feedback scores on eBay. Having estimated the model, I then run counterfactual simulations to find the optimal selling strategies.

I find that sellers are behaving as hard bargainers in the market. Namely, they do not want to deviate much from the BIN price they choose. Moreover, I also find that sellers’ behavior is optimal in the market.
The next section provide the details of the data. The parametric estimation strategy is described in Section 3.3. Results and discussion are provided in Section 3.4. Section 3.5 concludes.

3.2. Data

The data required for the econometric analysis was collected by a spider program coded in Java. From February 26, 2011 to May 16, 2011, I ran the spider periodically to collect data from BO listings for used Ford Mustang cars. For each car listed in this period, I have the information about the features of the car, listing properties, and seller characteristics. Moreover, I also collected book values of the cars from Kelley Blue Book (KBB) website\(^2\). KBB value captures the year, model, trim, transmission type, mileage of the car as well as the type of the seller i.e. dealer vs private seller. Nevertheless, it does not capture all the features of the car such as interior options including airbags, leather seats, cruise control etc. Thus, KBB value is only an approximation of the actual value of the car. For the empirical analysis, I focus on Mustangs that have clean titles and have years after 1990\(^3\).

As a result, I am left with 945 BO listings for which the summary statistics are given in Table 3.1. The average book value for the cars in the sample is 16,582 while the average BIN price is 21,207. Aside from listing and seller characteristics, eBay provides further information about cars. One of them is the Autocheck score of the car.

\(^2\)"www.kbb.com"
\(^3\)Cars with salvage or rebuilt title can be considered to have a different market. Also, KBB values are available only for years after 1990. Some cars could not be assigned a KBB value since they lack technical description. Furthermore, cars with an invalid vehicle identification number (VIN) or improper mileage such as 999,000 are also left out of the sample.
Table 3.1. Summary Statistics for BO Listings

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Mean</th>
<th>Std Deviation</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIN Price (in $)</td>
<td>21,207</td>
<td>11,856</td>
<td>1,400</td>
<td>59,995</td>
</tr>
<tr>
<td>KBB Value (in $)</td>
<td>16,582</td>
<td>8,483</td>
<td>1,675</td>
<td>37,643</td>
</tr>
<tr>
<td>Sale Price (in $)</td>
<td>17,002</td>
<td>11,361</td>
<td>1,250</td>
<td>56,500</td>
</tr>
<tr>
<td>Number of Offers</td>
<td>1.94</td>
<td>2.61</td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td>AutoCheck Score</td>
<td>80.55</td>
<td>19.93</td>
<td>2</td>
<td>96</td>
</tr>
<tr>
<td>Year</td>
<td>2004</td>
<td>4.95</td>
<td>1991</td>
<td>2011</td>
</tr>
<tr>
<td>Mileage</td>
<td>42,430</td>
<td>38,674</td>
<td>1</td>
<td>180,000</td>
</tr>
<tr>
<td>Seller Positive Feedback</td>
<td>51.17</td>
<td>209.52</td>
<td>0</td>
<td>4,280</td>
</tr>
<tr>
<td>Seller Negative Feedback</td>
<td>0.28</td>
<td>1.35</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>Warranty Dummy</td>
<td>0.25</td>
<td>0.43</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Dealer Dummy</td>
<td>0.50</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

car which is a number between 0 and 100, and grades the condition of the car checking all its registered history. Also, most of the cars in the sample have no warranty.

There are almost equal numbers of dealers and private sellers in the data. Furthermore, in the literature on online auctions seller reputation has been shown to be an important factor affecting the price (See Melnik and Alm, 2002; Houser and Wooders, 2006). After each transaction on eBay, the buyer is allowed to give the seller a feedback which can be positive, negative or neutral. I use the numbers of positive and negative feedbacks given in the last 12 months to control for the seller reputation.

As for the offer data, eBay discloses offer amounts when the listing ends. For each listing, I also observe the date and time that offers are submitted and the partially revealed IDs of buyers. Counteroffers by sellers are not observed even after the listing is completed. Hence, it is difficult to model the entire interaction between buyers and sellers in the absence of the counteroffer data. However, for the purposes of this study, focusing on single offers still allows me to keep the generality since single offers
Table 3.2. First Offers

<table>
<thead>
<tr>
<th>Ratio of Buyers Who Make Single Offers</th>
<th>63%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio of Buyers with only First Offers</td>
<td>74%</td>
</tr>
<tr>
<td>Ratio of First Offers among the Accepted Offers</td>
<td>73%</td>
</tr>
</tbody>
</table>

constitute the majority of the offer data as depicted in Table 3.2. Nevertheless, single offers are not actually identified in the data for sold cars.

This is because in listings that ended up with sale, one cannot know for sure whether seemingly single offers are actually single offers or just the first offers of buyers who could not have the chance to make another offer. Hence, such offers should carefully be treated in the empirical analysis. Excluding them from the data may result in selection bias since all accepted single offers would also be excluded in that case. In the empirical analysis, I estimated the model for both cases i.e. excluding and keeping the seemingly single offers. It turns out that the results are very similar for both cases as will be discussed more in Section 3.4. There is also a concern for listings that ended up with the BIN option. eBay does not reveal the identity of the buyer who exercises the BIN option. If there are previously made offers in these listings, one cannot know whether the buyer made an offer before choosing the BIN option. Nevertheless, the number of these listings is low (around 2%), and also to prevent a possible selection bias, I do include listings that ended up with the BIN option but did not receive any offer.

In my dataset around 63% of the buyers made a single offer, and 73% of the accepted offers are first offers. Summary statistics for 833 single offers in the data,

---

4The proportion of the seemingly single offers is around 15% of the data.
5This excludes the seemingly single offers. So this is actually a lower bound. 74% of the buyers only made the first offer. The actual ratio is between these two ratios.
Table 3.3. Single Offers

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std Deviation</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Offers (in $)</td>
<td>16,487.63</td>
<td>10,460.07</td>
<td>350</td>
<td>54,900</td>
</tr>
<tr>
<td>Number of Buyers*</td>
<td>2.25</td>
<td>1.6</td>
<td>1</td>
<td>12</td>
</tr>
</tbody>
</table>

*In listings with at least one single offer.

excluding the *seemingly* ones, are given in Table 3.3. The average single offer amount is 16,487 and lower than the average KBB value and BIN price. Also, on average 2.25 buyers participated in listings where a single offer is submitted.

3.3. Estimation

Estimation of the structural model is carried out in two steps. First, threshold distribution, \( G(.) \), is estimated. Then by using the estimates, \( \tilde{G}(.) \) and \( \tilde{g}(.) \), pseudo-values are obtained from equation (2.2), and \( F(.) \) is estimated in the second step. I will continue the analysis assuming a parametric form for \( F(.) \) due to the complications in nonparametric estimation mentioned in the previous chapter. I start the structural analysis with the first step estimation of \( G(.) \). Recall that Lemmas 1 to 3 are applicable under the assumption that \( G(.) \) is log-concave.

I assume that \( G(.) \) is a truncated normal distribution on the bounded interval between zero and the BIN price. In other words, for all \( t = 1, ..., T \), the lower bound is \( x_t = 0 \) and upper bound is \( x_t = p_t \). The selection of the upper bound is by definition of \( p_t \), and choosing the lower bound as zero is for convenience. If a positive lower bound is imposed, then the observed number of buyers differs from the number of potential buyers as in auctions with positive reservation prices. Though

---

6 Some buyers do not seem to behave rationally such as those submitting 1$. I left out such offers.

7 Note that the truncated normal distribution is log-concave. See Bagnoli and Bergstrom, (2005).
this can be added into the model, it also brings up some complications. First of all, there does not exist a rule-of-thumb for selecting a positive lower bound. So, imposing a prior lower bound would likely to induce bias unless there is empirical evidence justifying it. One can also consider estimating the lower bound but, as is well known, this is a problematic task due to the asymptotic behavior of bound estimators. Also, threshold prices are not observed in the data even after the listing ends, which complicates the issue further. So, I stay with the assumption that the lower bound is zero. Parameters of $G(.)$ are assumed to depend on covariates linearly.

Let $X_t$ denote the random variable for the unobserved threshold price for sale $t$ and $x_t$ be its realization. Then, $X_t \sim Trunc.Norm(\mu_t, \sigma_t)$ over the support $[0, p_t]$ where $\mu_t = \beta.Z_t$ and $\sigma_t = \alpha.Z_t$ are the mean and standard error of the normal distribution. $Z_t$ is the vector of covariates including a constant term and the KBB value of the car. $\beta$ and $\alpha$ are the corresponding parameter vectors.

To estimate the parameters $\beta$ and $\alpha$, I use the offer data which includes, for each listing, the offer amounts and whether the offers are accepted or rejected\(^8\). Since thresholds, $x_t$, are not observed in the data, estimation is carried out based on the acceptance-rejection decisions of sellers. Let $T$ be the number of listings in the sample and $\Omega$ be the set of listings. The offer data can be partitioned into three mutually exclusive sets for convenience: $A_1 = \{t \in \Omega :$ only rejected offers exist in sale $t\}$, $A_2 = \{t \in \Omega :$ both rejected offers and an accepted offer exist in sale $t\}$, and $A_3 = \{t \in \Omega :$ only an accepted offer exists in sale $t\}$. Because the threshold is assumed to

---

\(^8\)I assume that an offer is rejected if it is either explicitly declined by the seller or expired. Also, I use all the offer data (single and multiple offers) in the first stage estimation.
be fixed throughout the sale, only the largest rejected offer is needed when the number of rejected offers is more than one. By definition, only one accepted offer can exist in any sale. Since thresholds are unobserved, the likelihood function consists of the product of the probabilities for observing intervals. Let $b^a_t$ and $b^r_t$ be the accepted offer and the maximum rejected offer in sale $t$ (if they exist), respectively. Representing the entire offer data by $b = \{b^a_t, b^r_t\}_{t=1}^T$, and the covariate set by $Z = \{Z_t\}_{t=1}^T$, the log-likelihood function can then be constructed as follows:

$$
\log L_1(\alpha, \beta; b, Z) = \sum_{t=1}^T \left\{ I(t \in A_1) \log \left( \frac{\Phi_t(p_t) - \Phi_t(b^r_t)}{\Phi_t(b^a_t) - \Phi_t(0)} \right) + I(t \in A_2) \log \left( \frac{\Phi_t(b^a_t) - \Phi_t(0)}{\Phi_t(p_t) - \Phi_t(0)} \right) + I(t \in A_3) \log \left( \frac{\Phi_t(b^r_t) - \Phi_t(0)}{\Phi_t(p_t) - \Phi_t(0)} \right) \right\}
$$

where $\Phi_t(.)$ is the normal cdf with parameters $\mu_t$ and $\sigma_t$. The estimates for the parameters $\beta$ and $\alpha$ are obtained as,

$$
\{\hat{\beta}, \hat{\alpha}\} = \arg \max_{\alpha, \beta} \log L_1(\alpha, \beta; b, Z)
$$

Having estimated the threshold distributions and densities, $\hat{G}_t(.)$ and $\hat{g}_t(.)$, I can now plug them back into equation (2.2) together with single offers to get pseudo-values of buyers. Let $v_{it}$ denote the pseudo-value of buyer $i$ who submitted a single offer in sale $t$ which did not end up with sale. Also let $N_i$ be the set of these buyers in sale $t$, and $\Upsilon = \{ t \in \Omega : \text{There exists at least one single offer in listing } t \}$ and the car is
unsold\(^9\) be the set of such listings. Allowing for possible correlation across valuations, I assume \( f(v_{1t}, \ldots, v_{N_{tt}} | \mathbf{Z}_t) \) is a symmetric multivariate normal pdf conditional on car and seller characteristics with mean vector \( \overline{\mu}_t \) and covariance matrix \( \overline{\Sigma}_t \) such that,

\[
\begin{align*}
\overline{\mu}_t &= \left[ \theta \cdot \mathbf{Z}_t, \ldots, \theta \cdot \mathbf{Z}_t \right]_{1 \times N_t} \\
\overline{\Sigma}_t &= \begin{bmatrix}
\text{var}(v_{1t}) & \text{cov}(v_{1t}, v_{2t}) & \cdots & \text{cov}(v_{1t}, v_{N_{tt}}) \\
\text{cov}(v_{2t}, v_{1t}) & \text{var}(v_{2t}) & \cdots & \vdots \\
\vdots & \vdots & \ddots & \vdots \\
\text{cov}(v_{N_{tt}}, v_{1t}) & \cdots & \cdots & \text{var}(v_{N_{tt}})
\end{bmatrix}_{N_t \times N_t}
\end{align*}
\]

with, \( \text{var}(v_{it}) = \delta_1 \cdot \overline{\mathbf{Z}}_t \) and \( \text{cov}(v_{it}, v_{jt}) = \delta_2 \cdot \overline{\mathbf{Z}}_t \) for all \( i \neq j \) in listing \( t \).

The covariate set for listing \( t \), \( \mathbf{Z}_t \), includes a constant, the KBB value, the difference between the BIN price and the KBB value, autocheck score, seller positive and negative feedback ratings, warranty dummy (1 if it has). Also, \( \overline{\mathbf{Z}}_t \) includes a constant and the KBB value for listing \( t \). \( \theta, \delta_1, \text{and } \delta_2 \) are the corresponding parameter vectors to be estimated. The covariate set is denoted by \( \mathbf{Z} = \{ \mathbf{Z}_t, \overline{\mathbf{Z}}_t \}_{t=1}^{T} \) and the set of pseudo-values by \( \mathbf{v} = \{ v_{it} : i = 1, \ldots, N_t \text{ and } t \in T \} \). Furthermore, one also needs to consider the sales where the BIN option is exercised. A buyer chooses to pay the BIN price in sale \( t \) if her value is greater than the critical value i.e. \( v_{it} > v_c^t \) where \( v_c^t \) solves \( U_{O}(v_c^t; G_t) = U_B(v_c^t; G_t) \). Let \( B = \{ t \in \Omega : \text{In sale } t \text{ BIN price is exercised and there are no previous offers} \} \) be the set of such listings. The second stage log-likelihood function can then be constructed straightforwardly, and parameters are estimated as

\(^9\)I only focus on unsold listings because, as mentioned previously, single offers are not identified in sold listings. Moreover, the results are very similar when I repeat the estimation including the seemingly single offers in sold listings.
follows:

$$\log L_2(\theta, \delta_1, \delta_2; Z, v) = \sum_{t=1}^{T} \left\{ I(t \in \mathcal{Y}) \log f(v_{1t}, ..., v_{Nt}) | \mathcal{Z}_t, \mathcal{Z}_t \right\} + I(t \in B) \log \left( 1 - F(v^*_t | \mathcal{Z}_t, \mathcal{Z}_t) \right)$$

$$\{\hat{\theta}, \hat{\delta}_1, \hat{\delta}_2\} = \arg \max_{\theta, \delta_1, \delta_2} \log L_2(\theta, \delta_1, \delta_2; Z, v)$$

3.4. Results and Discussion

First stage estimation results are given in Table 3.4. 575 observations (intervals) are used in the estimation, and all parameters are significant. For a representative listing\textsuperscript{10}, the estimated threshold density is plotted in Figure 3.1. It is clear that the seller’s threshold density is highly concentrated close to the BIN price. As a result, given the BIN prices sellers act as hard bargainers in the market. In counterfactual simulations, I test whether setting such high thresholds is optimal for sellers or they should leave more room for bargaining.

<table>
<thead>
<tr>
<th>Mean</th>
<th>Std Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>KBB/1000</td>
</tr>
<tr>
<td>12.93\textsuperscript{**}</td>
<td>(6.44)</td>
</tr>
<tr>
<td>1.98\textsuperscript{***}</td>
<td>(0.58)</td>
</tr>
</tbody>
</table>

Standard errors are in parenthesis.
(*)\textsuperscript{,} (**), and (***) show significance at 10, 5 and 1\% levels, respectively.
Offers and BIN prices are in $1,000.

Plugging the estimates of the cdf and pdf of the threshold as well as the single offers in the FOC, I get the \textit{pseudo-values} of buyers. Using them for the second step of the

\textsuperscript{10}The KBB value and the BIN price for this listing are $15,480 and $17,888, respectively.
estimation, I get the parameter estimates in Table 3.5\footnote{Since estimation is carried out in two steps, I obtained the standard errors by bootstrap. To do this, I resample from the listings 1000 times.}. In all the specifications, the KBB value of the car affecting the mean parameter is significantly positive and close to one as expected. Though the KBB value seems to capture most of the variation in buyer valuations, there is almost always some information that cannot be picked up by the KBB value.

These may include other technical details of the car or some interior features of it such as music system, seat cover, airbags etc. If sellers are behaving rationally, then the BIN price they set should reflect such information as well. I use the difference between the BIN price and the KBB value as an additional information variable, and its significantly positive coefficient suggests that sellers indeed determine the BIN
Table 3.5. Structural Estimation Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Single Offers</th>
<th>First Offersa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Specific 1</td>
<td>Specific 2</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.11**</td>
<td>-0.0955**</td>
</tr>
<tr>
<td></td>
<td>(0.0515)</td>
<td>(0.0454)</td>
</tr>
<tr>
<td>KBB/10,000</td>
<td>0.8808***</td>
<td>0.9047***</td>
</tr>
<tr>
<td></td>
<td>(0.0261)</td>
<td>(0.0262)</td>
</tr>
<tr>
<td>(BIN-KBB)/10,000</td>
<td>1.0289***</td>
<td>1.0341***</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.0685)</td>
</tr>
<tr>
<td>Auto Check/100</td>
<td>0.1708**</td>
<td>0.142**</td>
</tr>
<tr>
<td></td>
<td>(0.0755)</td>
<td>(0.0645)</td>
</tr>
<tr>
<td>MeanSeller Positive Feedback/100</td>
<td>0.0045</td>
<td>-0.0003</td>
</tr>
<tr>
<td></td>
<td>(0.0077)</td>
<td>(0.0067)</td>
</tr>
<tr>
<td>Seller Negative Feedback</td>
<td>-0.009</td>
<td>-0.0025</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Warranty</td>
<td>0.0926**</td>
<td>0.0825*</td>
</tr>
<tr>
<td></td>
<td>(0.0466)</td>
<td>(0.0447)</td>
</tr>
<tr>
<td>((BIN-KBB)/10,000)²</td>
<td>-0.1884***</td>
<td>-0.1838***</td>
</tr>
<tr>
<td></td>
<td>(0.0332)</td>
<td>(0.0425)</td>
</tr>
<tr>
<td>Variance</td>
<td>-0.0026</td>
<td>-0.0044</td>
</tr>
<tr>
<td></td>
<td>(0.0169)</td>
<td>(0.0142)</td>
</tr>
<tr>
<td>KBB/10,000</td>
<td>0.1715***</td>
<td>0.1749***</td>
</tr>
<tr>
<td></td>
<td>(0.0179)</td>
<td>(0.0179)</td>
</tr>
<tr>
<td>Covariance</td>
<td>0.0198</td>
<td>0.0167</td>
</tr>
<tr>
<td></td>
<td>(0.0174)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>845</td>
<td>997</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-769</td>
<td>-885</td>
</tr>
</tbody>
</table>

Dependent variable is valuation or critical value in $10,000.

aThese include both single offers and the seemingly single offers.

price taking the quality features of the car into account, including those not captured by the KBB value.\(^{12}\)

An interesting result is about the effect of the seller reputation. As is well known, seller reputation has been shown to be an important determinant of price in online markets.\(^{13}\) Both of the reputation variables, positive and negative feedback ratings, are insignificant except for specification 2 where positive rating is weakly significant.

---

\(^{12}\)Note that BIN price is endogenously set by sellers, hence I do not give any causal interpretation to its coefficient. Moreover, in specification 2 of Table 5, I provide the estimation results excluding this information variable.

\(^{13}\)See, for example, Melnik and Alm (2002); Houser and Wooders (2006).
This may be due to the particular market I study here. Because sellers receive feedback regardless of what they sell\textsuperscript{14}, buyers in the car market may not view feedback ratings as a reliable measure for seller reputation.

In the last column of Table 3.5, I provide the results from estimating the model including the seemingly single offers. Recall that in the data observed single offers for sold cars may not always be what they seem. They may be the first offers of buyers who could not make another offer since the car is sold. Excluding such offers for sold cars may cause selection bias in the estimates given in specifications 1 and 2. To check for this, I repeat the estimation including the seemingly single offers and report the results in the last column of Table 3.5. Comparing them with Specification 1, I see that the estimates do not differ much. This implies that excluding the seemingly single offers for sold cars does not result in a problem.

Finally, it is worth discussing the estimates for the covariance matrix. The variance of buyer valuations significantly increases with the KBB value of the car. Both low and high valued buyers show interest in high valued cars. Furthermore, the covariance across buyer valuations is positive though it is weakly significant for the first specification. This contradicts the independence assumption commonly used in the literature on NYOP markets. The reason for this is clear. The covariates capture considerable amount of information contained in the listing but they do not cover all the information. The remaining information make buyer valuations correlated in

\textsuperscript{14}In other words, a positive feedback from selling a camera or car counts the same towards the feedback rating.
a listing. Also, buyers are able to communicate with sellers privately, and obtain additional information not posted in the listing.

Both of these facts create unobserved heterogeneity in the data which in turn results in correlation across valuations. When the additional information variable is excluded as in Specification 2, the effect of the unobserved heterogeneity gets larger as is reflected in the variance and covariance of the valuations.

Figure 3.2. Histogram of Actual Offers

Having estimated the structural model, I then check the goodness of the fit by simulating the model. Using the calibrated model in Specification 1, I simulate the BIN decisions and offers of buyers for each listing in the data. Histograms of the actual and simulated offers are shown in Figures 3.2 and 3.3. The estimated model seems to provide a reasonable fit to the offer data yet the BIN option is chosen more often in the simulated data. Thus, the number of simulated offers turns out to be
less than that in the data. The reason for this is the relatively high variance of the value distribution. On the other hand, the distribution of the simulated offers seems to match with the actual offer distribution in terms of skewness and weight.

3.4.1. Optimal Seller Strategy

One interesting aspect of the data is that sellers do not to leave much room for bargaining once they have chosen the BIN prices, though bargaining is one of the key features of the BO mechanism. The estimated threshold distribution shows that sellers only accept offers that are close to the BIN price they set. Using the calibrated model, I test the optimality of this seller behavior and other possible selling strategies allowing for different BIN prices and bargaining behavior.
First, I take the BIN price chosen by the seller as given and test the optimality of a hard bargaining strategy\textsuperscript{15}. To do this, I take the representative listing used in Figure 3.1 and simulate the buyer behavior under different bargaining strategies. Note that buyers do not observe seller’s threshold price, and take action only knowing its distribution. So given a commonly known threshold distribution, buyers cannot update their strategy in response to a one-time change in seller’s bargaining strategy since it is unobserved to buyers. However, buyers can observe the past sale data and notice a change in seller’s bargaining strategy that happened in the recent past. Having realized this change, buyers also update their strategies accordingly. In the simulations therefore, I follow an equilibrium type approach, and simulate the buyer behavior in an equilibrium that is reached after a period of interaction between the buyers and the seller following a change in the seller’s bargaining strategy. Taking the buyers’ responses into account, the seller determines her bargaining position to maximize her expected revenue. Although the model I use is decision-theoretic, I can still capture this interaction as in a game-theoretic model.

To simulate this environment for the representative listing, I change the bargaining strategy of the seller exogenously given the BIN price of the car. By shifting the mean of the estimated threshold distribution I try alternative bargaining strategies for the seller. More specifically, I simulate flexible bargaining by reducing the mean of the threshold distribution, and strict bargaining by doing the opposite. I then simulate the buyer and seller behavior using the threshold distribution for the

\textsuperscript{15}Being a hard or soft bargainer is, of course, relative to the BIN price the seller has chosen. I use this terminology for the seller behavior taking the seller’s BIN choice as given.
corresponding bargaining strategy. Buyers’ valuations are drawn jointly from the es-
timated multivariate normal distribution. I solve for the critical buyer value\textsuperscript{16}, \( v^c \), that makes buyers indifferent between paying the BIN price and making an offer. Given their private valuations, buyers play their optimal strategy. The seller draws a private threshold from the threshold distribution, and accepts or rejects offers accordingly. Repeating this 5,000 times I obtain the seller’s average revenue for each specified threshold distribution and for different reservation values of the seller\textsuperscript{17}.

As alternative bargaining strategies for the representative listing, I use the threshold densities given in the Figures 3.4 to 3.12. The mean of the estimated threshold density for the representative listing is denoted by \( \mu \). I construct the alternative threshold densities by assigning new mean parameters \( \mu' \). Densities with means greater than \( \mu \) represent stricter bargaining strategies for the seller. Whereas, densities having smaller means than \( \mu \) represent more flexible bargaining strategies given the BIN price.

For different reservation values, I simulated the seller’s expected revenue as in Figure 3.13\textsuperscript{18}. It is clear from the figure that the seller’s revenue is maximized when she becomes a hard bargainer given the BIN price she has chosen. This is consistent with what sellers do in BO sales. Leaving a little room for bargaining, sellers choose threshold prices close to the BIN price of the car to maximize their revenues. Moreover, choosing a threshold price is not the only decision sellers need to make.

\textsuperscript{16}I do this numerically using the FOC since closed form expressions are not available.

\textsuperscript{17}Reservation value of the seller is the amount the seller gets if the car is not sold.

\textsuperscript{18}The horizontal axes shows the multiples of the mean of the estimated threshold distribution, \( \mu \).
They first determine a BIN price for the car. In Figures 3.14 to 3.16\textsuperscript{19}, I also allow for different BIN price options in addition to the aforementioned bargaining strategies. For each BIN price-bargaining combination, I simulate the seller’s revenue 1000 times, and for each reservation value, I plot the expected revenue as a three-dimensional surface in Figures 3.14 to 3.16. The peak of each surface shows the revenue maximizing BIN price-bargaining strategy pairs for the seller. The BIN price-bargaining strategy chosen by the seller in the market turns out to be one of these strategies. Hence, I conclude that the observed seller behavior in the market is indeed optimal.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure3.4.png}
\caption{Bargaining Strategy with $\mu' = 3\mu$}
\end{figure}

\textsuperscript{19}The horizontal surface shows different combinations of BIN prices, and bargaining strategies as multiples of $\mu$. 
Figure 3.5. Bargaining Strategy with $\mu' = 2\mu$

Figure 3.6. Bargaining Strategy with $\mu' = \mu$

Figure 3.7. Bargaining Strategy with $\mu' = 0.5\mu$
Figure 3.8. Bargaining Strategy with $\mu' = 0.4\mu$

Figure 3.9. Bargaining Strategy with $\mu' = 0.35\mu$

Figure 3.10. Bargaining Strategy with $\mu' = 0.3\mu$
Figure 3.11. Bargaining Strategy with $\mu' = 0.25\mu$

Figure 3.12. Expected Seller Revenue vs Bargaining
Figure 3.13. Expected Seller Revenue vs BIN-Bargaining (Reservation Value = 0.8 * KBB Value)

Figure 3.14. Expected Seller Revenue vs BIN-Bargaining (Reservation Value = 0.6 * KBB Value)
3.5. Conclusion

In this study, I analyzed the "Best Offer" mechanism on eBay using a dataset of used Ford Mustangs. In particular, I estimate the buyer valuations using a structural model, and examine the optimal selling strategies. The model relaxes two standard assumptions commonly used in the empirical "Name Your Own Price" literature. First, I relax the independence assumption for buyer valuations to account for the unobserved heterogeneity in the data due to high information content and interactive features of listings. Second, I allow the seller to have a more general threshold distribution other than the uniform distribution.

In the econometric analysis, using the techniques from the literature on structural auction models I propose a two step estimation method. Estimation utilizes the first order condition, and does not require the computation of the offer function which
does not have an analytical representation. I then use the calibrated model to test
the optimality of the seller behavior in the market. Though the "Best Offer" format
creates a negotiation environment for buyers and sellers, sellers do not seem to deviate
much from the BIN prices they set. In other words, sellers behave as hard bargainers
in the market. I run counterfactual simulations to examine the optimality of this
behavior and other possible selling strategies. I find that sellers’ choice of BIN prices
followed by strict bargaining is indeed an optimal strategy for sellers.

Several other directions are available for future research in this area. It would
be very interesting to work with a more comprehensive dataset containing the entire
interaction between sellers and buyers. This would allow modeling multiple offers as
well as counteroffers in the market. It would also enable us to understand the reasons
that buyers use both single and multiple offer strategies. Moreover, I take each "Best
Offer" sale as an isolated environment. In general, on eBay there are many items sold
simultaneously or sequentially. Allowing for the intertemporal dynamics and search
behavior across listings could yield different implications. Focusing on the "Best
Offer" selling mechanism, which has not been analyzed much in the literature, this
study provides a step towards understanding it and opens a way for future research.
CHAPTER 4

An Incomplete Model for "Best Offer" Sales on eBay

4.1. Introduction

Measuring the willingness-to-pay or valuations of consumers has always been an interesting issue in the literature since knowing the willingness-to-pay (WTP hereafter), sellers can design profit maximizing selling strategies as well as launch new products. In this chapter, I propose a behavioral approach to estimate the distribution of WTP and its determinants for consumers who use the "Best Offer" selling mechanism on eBay which has not been studied much in the literature. I then apply the method to a dataset of used Ford Mustangs sold through the "Best Offer" format.

Different ways of measuring WTP have been proposed in the literature. In an open ended question format, consumers are asked to reveal their WTP directly, whereas in a choice-based conjoint analysis WTP is obtained from consumers’ choices among different products. In the absence of proper incentives, both methods can suffer from the hypothetical nature of the experiment. In an attempt to cope with this, Becker, DeGroot, and Marschak (1964) proposed an alternative mechanism where the consumer is to purchase the product if the price drawn from a lottery is less than or equal to her stated WTP. Miller et al. (2011), and Wertenbroch and Skiera (2002) compare the commonly used methods to measure WTP.
On the other hand, another stream of the literature focused on estimating WTP and its determinants using the real data. Naturally, empirical approaches to estimate WTP should be tailored for the specific selling mechanism used in the market. Due to their prevalence, auctions have garnered a lot of interest. Donald and Paarsch (1993) and Laffont et al. (1995) propose parametric estimation methods for structural auction models. Guerre et al. (2000) use nonparametric methods to estimate WTP. Bajari and Hortacsu (2003) use a common value auction model allowing for endogenous entry and utilizing Bayesian techniques. Chan et al. (2007) analyzes the effect of the market thickness and seller variables on WTP following the behavioral assumptions of Haile and Tamer (2003). Yao and Mela (2008) model both seller and buyer behavior in a structural model, and test how different strategies of the auction house affect the number of auctions and the bidding behavior.

Moreover, with the advent of the online markets new and hybrid selling formats have been launched, and researchers have considered methods to estimate WTP for these new selling mechanisms. Priceline has adopted the "Name your own price" mechanism where buyers make offers to buy flight tickets (See, for example, Wang et al. 2009, Terwiesch et al. 2005, and Amaldoss and Jain 2008). Spann et al (2004), and Hann and Terwiesch (2003) estimate WTP and cost of making offers for buyers in the "Name your own price" markets, respectively.

Similar to this mechanism, eBay launched the "Best Offer" format in 2005. Although other sale formats on eBay such as auctions and BIN sales have been studied extensively in the literature, the BO mechanism has not received the same attention (See Hasker and Sickles 2010 for a survey on eBay). To the best of my knowledge,
Chen et al. (2011) is the only other paper studying the BO format. Chen et al. (2011) study the BO environment in a sequential game setup and find reduced form evidence consistent with their model. However, they do not estimate the distribution of valuations.

In this study, I propose a way to estimate the distribution of willingness-to-pay of buyers who use the BO format. Following Haile and Tamer (2003), I use an incomplete model\(^1\) which is only based on two weak behavioral assumptions. There are two reasons for using an incomplete model. First, it allows one to capture many different equilibria that could arise in the "Best Offer" environment. Indeed, consistent with this fact, buyers stick to different strategies in BO sales. Some buyers make only one offer, whereas others offer multiple times. Since the subsequent offers of buyers depend possibly on the counteroffers by the seller and the counteroffer information is not available in the data, it would be complicated to model multiple offers in a complete model. However, one can capture multiple offers in an incomplete model which is the second advantage of using it. On the other hand, the downside of working with an incomplete model is that it does not allow the researcher to make counterfactual analysis.

Using a dataset of used Ford Mustangs, I focus on two main questions in this study: What are the determinants of WTP and how do they affect WTP? Are there any differences in WTP of buyers who offer only once and those who offer multiple times? To answer the first question, I apply the estimation method assuming that

\(^1\)I use the term "incomplete model" following Haile and Tamer (2003) to refer that buyer strategies are not fully specified.
buyers are homogeneous. I find that WTP is mostly explained by the book value of the car. For the second question, I estimate the model imposing heterogeneity for each group of buyers. I find that, for average and high value cars, the average WTP of buyers who offer multiple times is greater than that of buyers who offer once. The results suggest that sellers adopt group specific bargaining strategies. In particular, sellers could be better off following a harder bargaining strategy against multiple offerers.

The remainder of the chapter is organized as follows. The next section describes the econometric model. The results are given in Section 4.3. Section 4.4 concludes.

4.2. Econometric Model

Let \( T \) denote the total number of listings in the sample and \( N_t \) be the number of buyers interested in listing \( t \). Buyers are assumed to be risk neutral and draw their WTP independently from a continuous distribution, \( F(.)|X_t) \), conditional on the observed covariates. The covariate set, \( X_t \), includes the car and seller properties given in Table 3.1. Modeling the buying behavior I only assume the following two weak behavioral assumptions for the buyers:

\textit{Assumption 1: Buyers do not make offers more than their WTP.}

\textit{Assumption 2: If WTP of a buyer is higher than the BIN price, the buyer exercises the BIN option before the listing ends.}\(^2\)

\(^2\)The assumption implies that a buyer can follow any strategy during the listing; however, when the end time of the listing comes and if the buyer’s WTP is higher than the BIN price, then the buyer exercises the BIN option.
The assumptions are quite natural and allows me to stay agnostic about any equilibrium or optimality condition for buyers. They are also sufficient to capture multiple offers in the data which would, otherwise, be very complicated to account for in a complete model. Nonetheless, the cost of using such an incomplete model is the inability to make counterfactual experiments. In a complete model, one can obtain an offer strategy as a function of WTP, and use it to run counterfactuals to find optimal selling strategies for sellers. This is the trade-off involved in using an incomplete model.

I assume that $F(.|X_t)$ is a normal distribution with mean $\mu_t = \alpha \cdot X_t$ and standard deviation $\sigma_t = \beta \cdot X_t$ where $\alpha$ and $\beta$ are the vectors of parameters. Let $v_{it}$ and $b_{it}$ be the WTP and maximum offer of buyer $i$ in sale $t$, respectively. For any listing $t$ that ended up without sale, one can infer that WTP of buyer $i$ should be greater than her maximum offer and less than the BIN price, $p_t$, i.e. $v_{it} \in [b_{it}, p_t]$. On the other hand, if listing $t$ ended up with sale, then one cannot say more than the assumption i.e. $v_{it} > b_{it}$. This is because for sold listings, losers might not have had the chance to make another offer or exercise the BIN option since the listing ended prematurely, hence one cannot put an upper bound for their WTP. Similarly, though her offer is accepted, WTP of the winner cannot be bounded from above. Furthermore, I account for the listings that ended through the BIN option but did not get any offer\(^3\). Denoting the offer and covariate data by $b = \{b_{it}\}_{i,t}$ and $X = \{X_t\}_t$ for $i = 1, \ldots, N_t$

\(^3\)Recall that the identity of the buyer, who exercises the BIN option, is not observed in the data. Thus, if a BIN sold listing received any offer, one cannot use the winner’s information in the likelihood function since it could cause bias. However, offers in such listings are taken into account in the likelihood function.
and \( t = 1, \ldots, T \), the log-likelihood function can be constructed as follows:

\[
\log L(\alpha, \beta; b, X) = \sum_{t=1}^{T} \sum_{i=1}^{N_t} \{ I_1(t) \log (F(p_t|X_t) - F(b_{it}|X_t)) + I_2(t) \log (1 - F(b_{it}|X_t)) + I_3(t) \log (1 - F(p_t|X_t)) \}
\]

(4.1)

where, 

\[
I_1(t) = \begin{cases} 
1 & \text{if } t \text{ is an unsold listing} \\
0 & \text{otherwise}
\end{cases}
\]

\[
I_2(t) = \begin{cases} 
1 & \text{if } t \text{ is a sold listing with offers} \\
0 & \text{otherwise}
\end{cases}
\]

\[
I_3(t) = \begin{cases} 
1 & \text{if } t \text{ is a BIN sold listing without any offer} \\
0 & \text{otherwise}
\end{cases}
\]

The above estimation strategy treats all buyers symmetric and enables one to estimate the determinants of WTP for an average buyer. Nevertheless, the method can easily be extended to capture buyer heterogeneity by using buyer specific parameters. Because of the limitations of the data, I do not pursue that goal in this paper. Moreover, for the second research question regarding the difference between single and multiple offerers, I introduce heterogeneity for these two groups of buyers.

More specifically, I assume that WTP has a group specific normal distribution \( F_j(\cdot|X_t) \) where \( j = \{s, m\} \) differs for single and multiple offerers. The parameters of \( F_j(\cdot|X_t) \) are \( \mu_{jt} = \alpha_j \cdot X_t \) and \( \sigma_{jt} = \beta_j \cdot X_t \) for \( j = \{s, m\} \). Let \( v_{ijt} \) and \( b_{ijt} \) be the WTP and maximum offer of buyer \( i \) from group \( j \) in sale \( t \), respectively. The likelihood function is piecewise based on the group of buyers and whether a listing is sold or unsold. Because I am interested in analyzing WTP in different
groups i.e. single vs multiple offerers, I exclude buyers who exercised the BIN option. Furthermore, as mentioned previously, single offers are not identified in sold listings. Hence, I also exclude seemingly single offers in sold listings. As a result, the following likelihood function is maximized to estimate the group specific parameter vectors \( \{\alpha_s, \beta_s, \alpha_m, \beta_m\} \):

\[
\log L(\alpha_s, \beta_s, \alpha_m, \beta_m; b, X) = \sum_{t=1}^{T} \sum_{i=1}^{N_t} \sum_{j=(s,m)} \left[ I_1(t)D(i, j) \log (F_j(p_t|X_t) - F_j(b_{ijt}|X_t)) + I_2(t)D(i, j) \log (1 - F_j(b_{ijt}|X_t)) \right]
\] 

(4.2)

where, \( I_1(t) \) and \( I_2(t) \) are the same as before, and 

\[
D(i, j) = \begin{cases} 
1 & \text{if buyer } i \text{ is from group } j \\
0 & \text{otherwise}
\end{cases}
\]

**4.3. Results and Discussion**

As mentioned previously, an advantage of using an incomplete model is that it allows the researcher to use both single and multiple offers in the data. I estimate the distribution of WTP using the homogeneous model in equation (4.1) as well as the heterogeneous model in equation (4.2). Results are given in Table 4.1. As expected, the mean of the distribution of WTP is significantly related to KBB value. Similarly, Autocheck score has a positive effect on WTP though it is not significant in the heterogenous specification.

Furthermore, as is well known, seller reputation can be an important factor in online markets (Melnik and Alm, 2002; Houser and Wooders, 2006). However, both of the feedback variables used to capture seller reputation are insignificant in all
specifications. This could be due to the certain features of the feedback mechanism on eBay. The feedback ratings are aggregate in the sense that they reflect the total number of feedbacks received from selling anything on eBay. In other words, feedback ratings are not market specific, and a seller can receive a positive feedback after selling a car or a pair of socks. Therefore, buyers in the car market may not view the feedback ratings as reliable measures of seller reputation.

Table 4.1. Estimation Results

<table>
<thead>
<tr>
<th></th>
<th>Homogeneous</th>
<th>Heterogeneous</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Entire Data</td>
<td>Single Offers</td>
</tr>
<tr>
<td>Constant</td>
<td>0.2278***</td>
<td>0.2088***</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td>(0.0626)</td>
</tr>
<tr>
<td>KBB/10,000</td>
<td>0.9791***</td>
<td>0.9733***</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.0388)</td>
</tr>
<tr>
<td>Auto Check/100</td>
<td>0.1563*</td>
<td>0.0838</td>
</tr>
<tr>
<td></td>
<td>(0.0816)</td>
<td>(0.1061)</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seller Pos. Feedback/100</td>
<td>0.0055</td>
<td>0.0132</td>
</tr>
<tr>
<td></td>
<td>(0.0126)</td>
<td>(0.0154)</td>
</tr>
<tr>
<td>Seller Neg. Feedback</td>
<td>-0.0013</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>(0.0187)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>Warranty</td>
<td>0.0775</td>
<td>0.0364</td>
</tr>
<tr>
<td></td>
<td>(0.0629)</td>
<td>(0.0745)</td>
</tr>
<tr>
<td>Std Dev</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.2561***</td>
<td>0.2848***</td>
</tr>
<tr>
<td></td>
<td>(0.0272)</td>
<td>(0.0331)</td>
</tr>
<tr>
<td>KBB/10,000</td>
<td>0.2517***</td>
<td>0.2179***</td>
</tr>
<tr>
<td></td>
<td>(0.0205)</td>
<td>(0.0237)</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>1321</td>
<td>1157</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-2489.9</td>
<td>-2365.1</td>
</tr>
</tbody>
</table>

Regarding the second research question whether single offerers and multiple offerers have the same valuation structure, I estimate the model in equation (2) using only single and multiple offers. The results are given in the last two columns of Table 4.1. The parameters of the mean for each variable seem to be close to each other.
In order to check whether the differences between them are statistically significant, I conduct the following two sided test for each explanatory variable$^4$:

\[ H_0 : \text{Parameters for the variable are the same} \quad \text{vs} \quad H_a : \text{Parameters are different} \]

At 5% level, the null hypothesis is not rejected for all variables affecting the mean. I also do not reject the joint hypothesis whether all parameters are the same pairwise vs at least a pair of parameters is different. On the other hand, parameters of the standard deviation of WTP are found to be statistically different. Moreover, the mean and variance of WTP depend on covariates as well as parameters, and one would be interested in the difference between the mean WTP of single offerers and that of multiple offerers. However, one would want to know more than whether they are different. In particular, it would be more interesting from the managerial perspective to know whether the mean WTP of one buyer group is higher than that of the other group. Therefore, for each listing in the data, I use a one sided test with the following hypothesis:

\[ H_0 : \text{Mean WTP of single offerers is greater than or equal to that of multiple offerers} \quad \text{vs} \quad H_a : \text{Mean WTP of multiple offerers is greater than that of single offerers} \]

In 385 of 945 listings, the null hypothesis is rejected at 5% level (at 10% level, it is rejected in 613 listings). It turns out that most of these listings are for average or higher valued cars. In particular, the null hypothesis cannot be rejected for cars

$^4$As a result of maximum likelihood estimation, parameter estimates are asymptotically normally distributed. Based on this, the test statistic is also asymptotically normally distributed.
valued lower than $12,000 at 5% level. Moreover, it is rejected for around 63% of
average and high value cars (at 10% level, the rejection rate is around 94%).

On the other hand, comparing the variances of WTP of single and multiple offerers
in a similar test gives the following. At 5% level and for all cars valued higher than
$26,100 (the number of such cars is 125), the null hypothesis that the variance of WTP
of multiple offerers is smaller than that of single offerers is rejected. Also, for all cars
with values lower than $10,100 (the number of these cars is 285), the null hypothesis
that the variance of WTP of single offerers is smaller than that of multiple offerers
is rejected\(^1\). The results are illustrated in Figures 4.1 to 4.6 for three representative
cars.

\(^1\)Indeed, only for cars valued higher than $26,100 and lower than $10,100, the null hypotheses can
be rejected. Namely, the difference of the variances is not statistically significant for other cars.
Figure 4.2. cdf of WTP for a high value car (KBB = $27,344)

Figure 4.3. pdf of WTP for an average value car (KBB = $16,613)

Consistent with the test results, for the average and high value cars in Figures 4.1 and 4.3, the mean WTP is higher for multiple offerers than it is for single offerers. Also, the variance of WTP of multiple offerers is greater than that of single offerers for the high value car, and they are close to each other for the average value car. As for the low value car in Figure 4.5, however, the variance of WTP of single offerers become larger than its counterpart for multiple offerers, whereas the means are almost the same, which also parallels the test results. Furthermore, Figures 4.2, 4.4, and 4.6
Figure 4.4. cdf of WTP for an average value car (KBB = $16,613)

depict the cdfs of WTP for the corresponding cars. It turns out for the average and high value cars that WTP of multiple offerers has first order stochastic dominance over WTP of single offerers over a subset of the support that has most of the weight.

Figure 4.5. pdf of WTP for a low value car (KBB = $4,110)

To be able to interpret the results from a managerial perspective, one needs to characterize the seller side. In the literature on NYOP mechanisms, sellers are generally assumed to draw a private threshold, and only accept offers that are higher than their thresholds. The threshold price of a seller is expected to be a function of the
seller’s continuation or reservation value of not selling the item. Following the literature, I also assume this decision rule for BO sellers. Consider the sale of an average or a high value car. Assuming that the bargain price has exceeded the private threshold of a seller, the results suggest that the seller could extract more revenue from multiple offerers adopting a harder bargaining strategy against them. In other words, once a buyer makes the second or subsequent offers above the seller’s threshold, the seller could bargain harder knowing that the expected WTP of the buyer is likely to be higher than that of a single offerer. For cars with KBB values higher than $12,000 in the data, the average maximum and minimum offers of a multiple offerer is around $23,873 and $21,974, whereas the average single offer is around $21,136.\footnote{If all cars are included, the numbers are $16,810 and $15,316 for a multiple offerer and $16,488 for a single offerer, respectively.} It appears that the difference between WTP of single and multiple offerers are also reflected in their offers. Moreover, the difference between the maximum and minimum offers of

Figure 4.6. cdf of WTP for a low value car (KBB = $4,110)
a multiple offerer is also consistent with the suggestion that sellers bargain harder against multiple offerers. Sellers in the market seem to behave in the same direction.

4.4. Conclusion

In this study, I focus on a selling mechanism on eBay that has not garnered much attention in the literature. The BO format provides a bargaining environment for buyers and sellers allowing buyers to make offers below the BIN price, and sellers to respond these offers. Two research questions are studied in this chapter. First, I try to find out the determinants of WTP of buyers who use the BO format. To do this, I propose an incomplete model based on two weak behavioral assumptions. The second research question asks whether WTP of buyers who offers only once differs from WTP of buyers who make more than one offer.

Using a dataset of used Ford Mustangs sold through the BO format, I estimate the determinants of WTP. It turns out that the book value of the car is the major determinant of WTP, as expected. Seller reputation measured by the seller feedback ratings, however, becomes insignificant. One possible reason for this is that the feedback ratings are not market specific. In other words, sellers can receive feedback regardless of what they sell, and car buyers on eBay might not take the feedback ratings as accurate measures of seller reputation. Moreover, regarding the second research question, I find that buyers who offer more than once have on average a higher WTP than that of those offering once for average or high value cars. Given this finding, a differentiated bargaining behavior against each buyer group could be an optimal strategy for selling an average or high value car. Especially, sellers could
be better off bargaining harder with multiple offerers once the bargain amount reaches sellers’ reservation or threshold prices.

Several other directions are available for future research in this area. Because of the limitations in the data, I conduct the analysis in this paper for only different buyer groups i.e. multiple vs single offerers. However, as mentioned before, the econometric model can easily be extended to capture individual specific covariates. Hence, working with a more comprehensive dataset, one could come up with differentiated selling (or bargaining) strategies for each buyer. Furthermore, although the model proposed in this paper captures both single and multiple offers in the data, one cannot make counterfactual simulations because the model is incomplete i.e. buyer strategies are not fully specified. Using a complete model, one could make a more detailed analysis including counterfactuals to develop more specific optimal selling strategies.
References


