The Effects of Auction Parameters on Price Dispersion and Bidder Entry on eBay: A Conditional Logit Analysis

Eric Sun
esun@cs.stanford.edu
Department of Economics
Stanford University

Advisor: Professor Susan Athey
May 10, 2005

Abstract

This paper studies price dispersion and bidder entry in sequential online auctions of indistinguishable items, where goods and shipping costs are identical. Contrary to intuition and the “law of one price,” previous empirical analyses show large price swings among auctions of identical products. To study this phenomenon, this paper performs empirical analysis on a dataset of 3,164 sequential eBay auctions between May and June of 2004, all of which advertised invitations to beta-test Google’s invite-only Gmail email service. This novel dataset, unlike the auctions of coins and trading cards studied by previous papers, facilitates accurate analysis because there are no shipping costs and each auction supplies exactly the same product.

In spite of this obvious homogeneity, however, a huge spread was observed in the ending prices of these auctions—prices ranged from $0.99 to $217.50 per item. This paper studies existing theories of price formation and investigates how different parameters of an auction can influence its ending price. It also applies McFadden’s choice model, which uses conditional logit regressions, as part of an in-depth study of the effects of auction characteristics on bidder entry. Since this paper is the first to apply McFadden’s model to online auctions, it significantly extends existing literature on bidder choice. The paper then uses these results to develop a general theory of bidder behavior for private-value auctions on eBay. Finally, to conclude the study of price dispersion in sequential auctions, the paper empirically verifies the predictions of Weber’s martingale theorem; that is, previously ended auctions will influence the ending price of future auctions.

Keywords: auction, reserve price, martingale, competition between bidders, eBay, price dispersion, McFadden’s choice model, conditional logit

Acknowledgements: I would like to thank my advisor, Professor Susan Athey, for offering insightful suggestions throughout the course of my research. Special thanks also to Philo Davidson, Dr. Mary Sprague, Dr. Geoffrey Rothwell, and my family and friends for their support.
Table of Contents

1. INTRODUCTION ...........................................................................................................3
2. AUCTION BACKGROUND ..................................................................................................6
3. PAST AUCTION LITERATURE ..............................................................................................8
4. EMPIRICAL QUESTIONS .......................................................................................................18
5. MODELS ..............................................................................................................................18
6. EMPIRICAL APPROACH ......................................................................................................23
7. EMPIRICAL DESIGN ............................................................................................................26
8. METHODOLOGY AND RESULTS .........................................................................................32
9. CONCLUSION ......................................................................................................................50
APPENDIX A: SAMPLE AUCTION PAGE ...................................................................................53
APPENDIX B: CORRESPONDING BID HISTORY PAGE ...............................................................54
APPENDIX C: AUCTIONED PRODUCT .....................................................................................55
APPENDIX D: SOURCE CODE ..................................................................................................56
APPENDIX E: STATA CODE ......................................................................................................64
REFERENCES ..........................................................................................................................65
1. Introduction

Although auctions have been used as a market mechanism for thousands of years,\(^1\) they have historically been limited in scope to specialized arenas and thus have been foreign to the general public. For example, in the past few decades, government regulators have used auctions to sell telecommunications spectrum licenses and oil leases. With billions of dollars worth of revenue at stake, finding an optimal strategy is clearly valuable. Many academic papers have been published to deal with utility maximization for both buyers and sellers in such auctions, but the low number of potential bidders limits the effectiveness of empirical study. Recently, however, online auction marketplaces such as eBay have come into widespread everyday use. With eBay, any listed item can immediately capture the attention of millions of bidders, yielding a perfect framework for further experimentation.

There is much existing literature on the theory of auction pricing. Standard models use the revenue equivalence theorem (Myerson 1981), which finds that if two auction formats award a single object to the same bidder and exclude the same set of bidders, then they must yield the same expected revenue to the seller. In practice, however, things are not so simple. Since all bidders have different patience levels and use different bidding strategies, it is impossible to satisfy Myerson’s assumptions. In addition, analysis is complicated because sellers are allowed to specify dozens of options and parameters for each auction listing (e.g., time length, reserve price, and ending time), and each of these parameters has a measurable effect on an auction’s final ending price. This paper will use real auction data collected from eBay to quantify the effects of these variables on both auction price and bidder choice decisions. Though auction pricing has been extensively studied, bidder entry has not. This paper will contribute to the

\(^1\) The Greek historian Herodotus reported the use of Dutch auctions for selling brides in Babylonia.
literature on bidder entry models and use empirical results to motivate a general theory of bidder behavior on eBay.

Understanding bidder behavior is important for many reasons. If sellers can predict bidder entry patterns, they can modify their auction parameters to generate the highest possible revenue. Expert bidders may adjust their strategies to account for the actions of their counterparts. Even the auctioneer, eBay, may use this knowledge to fine-tune its revenue collection methods. For example, if eBay wanted to encourage use of their proxy bidding system and discourage naïve incremental bidding strategies, they could charge bidders an extra fee for each bid placed.

Previous empirical work on eBay has focused on datasets of items such as collectible coins, trading cards, and IBM ThinkPad laptops. However, there are several problems with these datasets: first, item condition is a critical issue. Since used items will sell for a lower price than new items, qualitative characteristics (such as graphical and text descriptions, where the item’s condition is described) influence prices greatly. This introduces noise in the data that cannot be easily corrected with quantitative methods. Another problem is that shipping costs must be taken into account. Hossain and Morgan (2003) have noted that bidders are not sensitive to shipping charges as long as they fall in some reasonable range, so biases can result if this is phenomenon not addressed. Finally, it is difficult to analyze auctions that generate no bids at all because the definition of “auction ending price” becomes ambiguous.

The novel dataset used in this paper solves each of these problems. I study 3,164 sequential eBay auctions that ended between May and June of 2004; all of these advertised invitations to beta-test Google’s invite-only Gmail email service. This is a much better dataset for empirical work than the ones used by previous authors because there are no shipping costs
and each auction supplies exactly the same product (consequently, the item’s condition is not a concern). In addition, item descriptions and/or pictures included on the auction page do not give bidders any additional information, except perhaps a signal of a seller’s good faith. Finally, because of the hype surrounding Gmail’s launch, only a handful of the auctions in the dataset ended with no bidders. Therefore, these auctions can be safely ignored in the empirical analysis. Eliminating these difficult-to-measure variables reduces much of the noise found in traditional empirical analysis of eBay data.

The Data

Although the Gmail system was announced to great public fanfare, actually signing up for an account was nontrivial due to the lack of a public signup page. Rather, an invitation system was set up to slowly increase the size of the program at Google’s desired pace. At first, accounts were generally limited to Google employees and included a specified number of invitations (one such invitation is displayed in Appendix C). Employees were free to invite their friends to try Gmail, and these new users were soon given their own invitations to give out as well. Not surprisingly, a market soon flourished for these invitations on eBay, with hundreds of invitations per day selling from anywhere between $0.99 and $217.50 per invitation. As an interesting side note, reselling of Gmail invitations was not expressly forbidden in the terms of service until several months after the start of the beta test program. Rumors have circulated of at least one Google employee being fired for attempting to resell Gmail invitations.
2. Auction Background

Before discussing previous literature on the subject, I will describe the key characteristics of auctions on eBay since these attributes will critically influence the empirical results of this paper. Though classic auction literature can be applied to online auctions as well, there has been a wealth of research that deals with the features of online auctions in particular. In addition, many tendencies of bidder entry have arisen from the rules and framework of eBay (for example, the practice of “bid sniping,” to be discussed in depth later, is impossible on auction sites like Amazon.com Auctions, where variable auction end times are used). Therefore, understanding various features of the eBay platform is necessary to ensure the clarity of the following sections.

eBay’s Feedback System

The eBay reputation system is based on a numerical system of feedback from a user’s past auction history. Although the rules for leaving feedback have changed over the years, the basic premise remains the same—after each transaction, the buyer and seller each have a chance to rate the other party. If a user has a positive experience, she will generally leave positive feedback for the other party; the opposite is also true for negative experiences. Each time a distinct user issues positive feedback, the recipient’s feedback score will go up by one point. Similarly, a negative comment reduces the recipient’s feedback score by one point. Although the system is entirely voluntary, users generally contribute to the eBay community by leaving feedback—indeed, it is not uncommon for a longtime seller to have a feedback rating in the tens of thousands. The feedback system is important because it allows users to gauge the possibility of default: that is, the chance that a winning bidder will not pay for an item or the chance that a
sells will not deliver the item after receiving payment. A user’s reputation can also be used as a proxy variable for her relative experience, since longtime users of eBay will clearly have larger feedback scores than novice users. For these reasons, feedback ratings of both buyers and sellers will be used as variables in later regressions.

*Buying and Selling on eBay*

To list an item on eBay, a seller must first select an auction category. For the dataset modeled in this paper, the category *Computers & Networking > Software > Internet Related Utilities > E-Mail Software* is generally chosen. Then, several features of the auction must be selected and published as public information. The most relevant choices are the auction duration (i.e. 3, 5, 7, or 10 days), the minimum bid, item description, and if desired, a secret reserve price. If the final price of the auction does not exceed the reserve price, the sale will not proceed. The seller can also choose the type of auction: it can be a regular English-style second-price ascending auction, a fixed-price ("Buy it Now!") auction, or a second-price auction with a "Buy it Now!" designation that disappears when the first bid is placed. Multi-unit auctions may also be created; however, for reasons to be explained later, these auctions will not be analyzed in this paper.

After the seller submits this information, eBay will post the auction to the public along with the seller’s reputation data. Prospective buyers can now search for, view, and bid on the auction. They can bid in several different ways—one possibility is to execute a naïve bidding strategy where they continually submit the minimum allowable bid until they are the highest bidder. Alternatively, eBay provides a powerful tool called proxy bidding. With proxy bidding, a buyer can enter the maximum amount that she is willing to bid. eBay will then submit bids on
the buyer’s behalf as the auction proceeds, bidding only enough to outbid the other bidders (taking into account eBay’s minimum bid increment). This process continues until the buyer’s maximum bid is exceeded, the auction ends, or the auction is won. eBay encourages all bidders to use the proxy bidding system and bid their reservation price early in the auction, claiming that this would constitute a dominant strategy. Their argument is that even if the bidder loses, she should not care because the winning bid will be above their reservation price anyway. In actuality, this is not strictly correct because of the time factors inherent in the online bidding process—since bidders become aware of the auction at different times, other bidding strategies (e.g., late bidding) may actually deliver higher expected utility for the buyer. These strategies will be discussed in the next section.

3. Past Auction Literature

Although the fervor surrounding online auctions has intensified only in the past few years, auction literature has stretched back for decades. Since William Vickrey’s seminal work “Counterspeculation, Auctions, and Competitive Sealed Tenders” was published in 1961, auctions have been studied extensively. In particular, Paul Milgrom and Robert Weber were influential in extending the theory of auctions, studying the revenue properties of many different types of auction formats. In recent years, with the rise in popularity of online auctions—millions of items are now listed every day on eBay—a similar boon has occurred in auction literature, much of it concerning pricing issues of online auctions in particular. These studies extend the work of the pioneers of the field by attempting to predict closing prices of online auctions; they study factors that contribute to bidding, such as the dynamics of private vs. common value
auctions, reputation models, and the phenomenon of late bidding (also known as “bid sniping”). Though these factors are important determinants of auction prices, there has been little investigation into how these parameters influence bidder choice. After discussing prior work in the field, I will expand the literature by introducing a model for bidder entry.

The Framework of Auctions

Different types of auctions deliver different expected revenues for the seller. Although the goal of any auction is to help buyers and sellers come to an agreement despite their unwillingness to reveal private information, different auction formats will generate different outcomes. Milgrom and Weber’s well-known work on auction theory (1982) gives a ranking of revenues across different auction rules in common values environments. Among their noted findings is that second-price sealed-bid auctions will provide higher seller revenues than first-price sealed-bid auctions, and a special kind of Japanese ascending auction yields even higher revenues than the second-price auction.

With all these auction mechanisms, finding the best auction format seems like a complex task since a strategic equilibrium and optimal strategies must be found for every type of auction. However, Myerson’s revenue equivalence theorem (1981) can be applied, at least in theory—his “revelation principle” states that if we have truthful reports of bidders’ valuations, then two auction mechanisms that award the good to the bidder with the highest valuation and deliver zero payoff for losers must yield identical ex-ante expected revenues for the seller.

Myerson’s lemma is a result closely related to the revenue equivalence theorem. Here, we consider a standard independent private values model. Expected payoffs of a Bayes-Nash equilibrium satisfy
\begin{equation}
V^i(\tau) = V^i(0) + \int_0^\tau E^i[z^i(\tilde{r} | t^i = s)] \cdot \frac{dv^i}{ds} \, ds,
\end{equation}

where \( t \) is the buyer’s type and \( V^i(\tilde{r}^i) \) is the maximum expected payoff of player \( i \) of type \( \tilde{r}^i \). If \( V^i \) is differentiable at \( \tau \), then

\begin{equation}
\frac{\partial}{\partial \tau} V^i(\tau) = E^i[z^i(\tilde{r} | t^i = \tau)] \cdot \frac{dv^i}{d \tau}.
\end{equation}

Expected payments must now satisfy

\begin{equation}
E^i[p^i(\tilde{r} | t^i = \tau)] = -V^i(0) + E^i[z^i(\tilde{r} | t^i = \tau)] \cdot v^i(\tau) - \int_0^\tau E^i[z^i(\tilde{r} | t^i = \tau)] \cdot \frac{dv^i}{ds} \, ds.
\end{equation}

See Milgrom (Milgrom 2004, p.75) for a proof of this result.

Unfortunately, eBay auctions in practice do not award the good to the bidder with the highest valuation because of time-varying parameters such as the “Buy it Now!” option. Thus, the revenue equivalence theorem alone is insufficient to predict the ending prices of auctions in practice.

In fact, empirical researchers have found that various parameters can significantly influence the auction’s ending price. For example, Lucking-Reiley, Bryan, Prasad, and Reeves (2000) show that auctions with a long time span generally end with higher prices than auctions with a short time span. They found that 7-day auction prices were approximately 24% higher than 3-day and 5-day auctions of the same product, and 10-day auctions were 42% higher. This makes intuitive sense, as longer auctions will be seen by more potential bidders and thus should receive more bids. A secondary conclusion of their paper is that minimum bids and reserve prices tend to have positive effects on the final auction price, though it also decreases the probability that the auction will result in an actual sale. I will use the work of these authors and include dummy variables for auction duration in my models of bidder entry.
The Role of Reputation

A number of economists have attempted to quantify the value of reputation in eBay auctions. Reputation is important in eBay auctions because there is a positive probability that a buyer will not deliver payment if she wins the auction, and there is similarly a positive probability that the seller will not deliver the item after receiving payment. However, the reputation system is not perfect since eBay does not require the two parties to leave feedback after a sale. In fact, there are several reasons why a user may not do so. Some novice users may not be aware of the system or will simply not take the time to leave feedback due to lack of incentive. Others will rationally choose not to. McDonald and Slawson (2002) show that users have an incentive not to provide negative feedback when appropriate due to the risk of retaliation. If a user feels that the other party will retaliate by issuing negative feedback in return, the user may instead choose not to leave feedback at all. Thus, eBay’s reputation system reveals only part of the private information of its users.

Although not without flaws, the eBay reputation system still provides an adequate signal for buyers and sellers. Several studies have performed regressions to estimate how much prices increase or decrease as a function of a user’s reputation. Lucking-Reiley, Bryan, Prasad, and Reeves studied a dataset of 20,000 auctions of collectible one-cent coins, finding that a 1% increase in the seller’s positive feedback ratings yields a 0.03% average increase in the final auction price. The effect of negative feedback ratings is much larger; a 1% increase causes a 0.11% decrease in auction price. The paper finds that the effect of negative feedback is statistically significant at the 5% level, but the effect of positive feedback is not.

Houser and Wooders (2000) find similar evidence using a dataset consisting of auctions of Pentium III 500 Mhz processors, finding that a 10% increase in positive feedback will
increase the winning bid by 0.17% and a 10% increase in negative feedback reduces the sale price by 0.24%. They also find that the effects of buyer reputation on price are statistically insignificant. Although this paper agrees with the authors’ conclusion that prices are correlated with seller reputation, I also find statistically significant evidence that buyer reputation influences the ending price as well. This could happen if experienced bidders have lower search costs, resulting in increased mobility among competing auctions. This mobility allows the bidders to find a better deal than they would have otherwise. A model for such a scenario will be proposed later in this paper.

Late Bidding

It has been observed that in online auctions of all types, the vast majority of bids occur in the very last stages of the auction. Bajari and Hortaçsu’s survey on coin auctions reports that “the median winning bid is submitted after 98.3% of the total auction has elapsed (the last 73 minutes of a 3 day auction) and 25% of the winning bids arrive after 99.8% of the auction time has elapsed (the last 8 minutes of a 3 day auction)” (Bajari and Hortaçsu 2002, p.12). Many papers have been written on the subject, each positing theories and models to justify bidders’ actions. The general consensus is that in hard close auctions (as opposed to auctions that continue until a specified time interval has passed with no further bids, as studied by Ariely, Ockenfels, and Roth (2002)) late bidding (also known as “bid sniping”) is a rational strategy, especially in auctions with common values.

In common values auctions, bidders can often be uncertain about the true value of the object that they are bidding on. For example, in Bajari and Hortaçsu’s coin auctions, a large fraction of the items are collectibles, so the quality of the object matters. In such auctions,
bidders can obtain information from others that may cause them to revise their willingness to pay. Late bidding is rational here because if A and B have their own private estimate, $x_i (i = A, B)$, about the object’s worth, then the average of their estimates, \( \frac{1}{2}(x_A + x_B) \), will be a better estimate of the object’s worth than either A or B’s estimate alone. Thus, it is advantageous to be the second bidder. In equilibrium, this results in both A and B bidding as late as possible. For this story to be true, it must be the case that bidders would change their own estimate of the object’s worth if they knew the estimates of the other bidders.

Roth and Ockenfels (2002) also point out that if a known expert is seen to have bid in an auction (since eBay bid histories and users’ reputations are public information, a bidder’s “expert” status can be easily ascertained), this will greatly increase the bid prices in a common values auction because the auction will in effect have been authenticated for free. Thus, late bidding enables bidders to avoid sharing valuable information with other bidders.

In the case of online auctions of Gmail accounts, the quality of the invitation is not an issue since every invitation is superficially identical. Because of this homogeneity, bidders will generally gain no information from seeing the identities of other bidders. In addition, although there may be a common-values element (because bidders will value an invitation more if they know many other people also value it), this effect will usually be small because of the lack of a secondary market for Gmail invitations. There is very little resale value for an account because invitations automatically expire if the invitee does not respond quickly enough. More importantly, the invitations are personalized with the new user’s name and alternate email address. Thus, it is generally frowned upon to resell second-hand Gmail invitations, and any common-value element to Gmail account auctions is assumed to be insignificant.
Even in such auctions with private values, however, bid sniping can be a rational strategy. Roth and Ockenfels (2002) model late bidding as a collusive attempt among bidders to keep prices low. By not entering the bidding until the very final stages of the game, the bidders avoid a prolonged bidding war with incremental bidders. Thus, assuming bidders care only about the current auction price, it is a Nash equilibrium for all bidders to bid snipe.

Schindler (2003) presents alternative theories that account for illicit motives on the part of auctioneers and sellers. She points out that because an auctioneer (i.e., eBay) takes a percentage of each sale as commission, it has an incentive to increase the ending prices of auctions by automatically using the highest bidder’s proxy bid instead of following the incremental bidding strategy they advertise. This moral hazard scenario is unlikely, but submitting a large proxy bid will not be a dominant strategy for such paranoid bidders. A far more common illegal practice is the procedure of shill bidding, where the seller bids on her own auction in order to inflate prices. By bidding late, a buyer can help ensure that no such activity has occurred because the seller will have less time to engage in such illegal activities. Indeed, in late 2004 a group of bidders were successfully prosecuted for engaging in shill bidding on eBay,² so can be a very real concern.

This paper acknowledges the existence of bid sniping and explains why bidders choose to use this strategy. In addition, I explore other tendencies employed by bid snipers and explain why these actions are rational.

---

² In November 2004, New York attorney general Eliot Spitzer announced civil settlements of up to $50,000 in several distinct cases of shill bidding on eBay, thus proving that the practice is indeed real and more widespread than commonly believed.
Minimum Bids and Reserve Prices

By setting minimum bids and reserve prices, sellers can adjust the amount of risk they want to take. If buyers always bid their valuations, sellers should maximize efficiency by setting the minimum bid to their own valuation of the object. This way, they are guaranteed not to lose money on the sale and will make money if potential buyers are competing for the item being sold. Bajari and Hortacsu (2002) note that in a second-price ascending auction, the minimum bid binds when it is between the highest and second-highest bids. If the minimum bid is within the bidders’ valuation distribution, it binds with positive probability. Thus, even if the seller derives zero utility from keeping the object, there is an incentive to set a positive minimum bid and prevent some efficient trades from occurring. Setting a high minimum bid ensures that the seller will not end up with an unexpectedly low ending price, assuming that at least one bidder participates in the auction.

However, under risk neutrality this strategy may not be in the seller’s best interest. An interesting question arises because there may also be a reason to set a low minimum bid to attract a larger audience, especially when demand for a good is known to be very high. McAfee, Quan, and Vincent (2002) propose an endogenous-entry model where sellers choose a reserve price and announce their auction to N potential bidders. These bidders then make a mixed-strategy entry decision to decide whether to incur participation costs. Finally, the bidders submit their bids after finding out private information about the value of the good. The authors assert that the optimal reserve price is at least as high as the salvage (alternate-use) value for the good.

Empirical evidence for this is mixed. While McAfee’s real estate auction data supports his theory, others find the opposite conclusion. Lucking-Reiley’s experiment using auctions of Magic cards on an Internet newsgroup concludes that low-reserve auctions generally end with a
higher price than auctions with high reserves. Lucking-Reiley explains this phenomenon by positing that there must be positive search costs, so buyers tend to stick with the auction in which they submit their first bid. This paper tests for this effect and finds statistically insignificant results, implying that these theories may both be true for different types of bidders.

*Sequential Auctions*

Several researchers have studied a special case of auctions where several identical items are sold in succession. This model provides a close parallel to the dataset described in this paper, assuming that the ending time between successive eBay auctions of Gmail accounts is relatively small (this is generally true). Weber’s martingale theorem (1983) shows that if the prices are announced after each sale, then the sequence of prices forms a martingale. This means that the expectation of the \( n + 1 \)” price given the prior prices is equal to the \( n \)th price. This model will be discussed in depth later in this paper.

Other economists have attempted to test for martingales in real-life auctions. Ashenfelter (1989) studies a sequence of wine auctions and finds clear evidence of a declining pattern in the prices of these sequential items, contrary to the prediction of Weber’s theorem. He observes that in a series of auctions for 1961 Chateau Palmer wines, the first lot went for 920 pounds; the second lot for 800 pounds; and the third lot for 700 pounds. Clearly, as the first bidder paid 31% more than the third bidder for the same item, the “law of one price” did not hold here. He finds that in the long run, prices are twice as likely to decrease as to increase in auctions of sequential items. Black and deMeza (1992) explain this phenomenon by noting that these declining prices exist in Ashenfelter’s wine auction because the winner of the first auction had the option to buy the remaining objects at the winning price. Yet the phenomenon has been empirically shown to
exist even when the option is not permitted, such as with transponder lease auctions in Milgrom and Weber (1982b).

Many other theories have suggested reasons for the anomaly; several are relevant for this paper. McAfee and Vincent (1993) show that risk aversion can create declining prices. Von der Fehr (1994) shows that declining prices can arise from participation costs. Pezanis-Christou (2001) explains the price decline to heterogeneity among buyers, and Ginsburgh (1998) shows that the presence of absentee bidders can generate declining prices.

Several authors have also found increasing prices in sequential auctions. Milgrom and Weber (1982b) show that if bidders’ valuations are affiliated, then prices will tend to rise over time in a sequence of identical objects. Empirical evidence for this is provided by Gandal (1997) for Israeli cable television licenses and Donald, Paarsch, and Robert (1997) for Siberian timber-export permits. This paper will describe Weber’s theoretical model and empirically verify its predictions.

Conclusion: Lessons for an Analysis of eBay

This section has described many important phenomena that have been shown to occur in auctions in practice. Because they can significantly influence the ending price of auctions, the following sections must take these factors into account as a more complex model of bidder entry is advanced. Although most of the papers I have discussed were published long before eBay’s heyday, their results are still very relevant. I will use their results to determine baseline variables used in the empirical analysis and then extend the models from there. For example, since previous authors have noted that parameters such as auction duration, seller reputation, and reserve prices have a statistically significant effect on auction results, I will include variables to
account for each of these effects. In addition, the discussion on late bidding motivates an analysis of whether certain types of bidders are more likely to bid differently than others. I will extend the work of previous authors and develop a general theory of bidder entry on eBay.

4. Empirical Questions

In my analysis of the Gmail auctions dataset, I attempt to shed light on the following questions:

1. How do the parameters of an auction influence the final ending price?

2. Specifically, does an auction with a reserve price deliver lower or higher returns for the seller when there are a large number of experienced and inexperienced bidders?

3. Is there empirical evidence that experienced bidders have lower entry costs than inexperienced bidders?

4. Is there evidence of price dispersion in sequences of auctions? That is, does one auction’s ending price influence the ending price of the next auction that closes?

5. Models

The effects of standard auction parameters have been studied in great detail by other researchers. However, as mentioned earlier, the datasets these authors use may be biased due to unobserved quality issues—the condition of an item is an important omitted variable that may skew results. Thus, tests of Empirical Question #1 will verify previous authors’ results in a setting where unobserved quality is not a concern. For the other questions, however, there is no
obvious conclusion. To answer Empirical Question #2, we must decide whether setting a low or high reserve price will deliver higher returns to the seller. If a seller sets a low reserve, the risk is undeniably higher—it is always possible that the auction somehow falls through the cracks and generates very little bidder activity. In addition, low-reserve auctions generally have longer durations, which is a problem if prices tend to fall over time. Thus, setting a high reserve may be better for a risk-averse seller.

However, in situations with high levels of bidder competition, setting a low reserve can actually give higher average returns if entry costs exist (Empirical Question #3): if buyers are more inclined to stick with the auction in which they place their first bid, then buyers who are more mobile will tend to profit. Thus, sellers want to increase competition and attract as many different bidders as possible in order to drive up the auction price. The following model illustrates this idea.

Model 1: (Entry Costs)

**Definition 1:** There are two types of bidders, inexperienced (denoted \( n_L \)) and experienced (denoted \( n_H \)). Bidders are segmented into the two categories under the requirement that \( n_L \approx n_H \) and that each inexperienced bidder has a feedback rating lower than or equal to all experienced bidders. Furthermore, assume that \( n_L + n_H \) is sufficiently large such that even without a reserve price, it is not an equilibrium for all bidders to bid in the auction due to eBay’s minimum bid increment rules.

Each bidder has singleton demand and wishes to win one of \( x \) identical auctions. The bidding process is performed as follows: each bidder searches eBay for a text string (i.e. “Gmail”) and is returned a listing of auctions with limited information: auction title, current
price, number of bids, and time left. Now, after seeing the choice set of auctions, a bidder must pay a fixed entry cost to click through and evaluate one auction in particular.

**Definition 2**: Denote the switching costs by $f_L$ and $f_H$ for inexperienced and experienced bidders. These costs must be incurred if and only if bidders bid in a different auction than the auction they first bid in.

**Proposition**: Assume that there are a large number of both low-reserve auctions and high-reserve auctions with minimum bid increment $\varepsilon$. Furthermore, assume that $f_L > f_H$ and that there are a large number, $n$, of potential bidders with identical values that bid incrementally in the lowest-price auction; that is, if the current high bid in one auction plus $\varepsilon$ is less than their value for the object, they will place a bid. Bidders have singleton demand, so they will not be the standing high bidder in more than one auction. Therefore, if the current high bid plus $\varepsilon$ is below their value for both auctions, they will bid in the cheaper one. Now, the high-reserve auction will end at a price $f_H$ below the price of the low-reserve auction, and the experienced bidder will win the lower-priced auction.

**Example**: Assume that there are two auctions. Auction A has a reserve price of zero, and auction B has a reserve price of 50. Assume that $f_L = 10$, $f_H = 5$, and $\varepsilon = 1$. Furthermore, assume that there is one inexperienced bidder and one experienced bidder.

Both bidders will start bidding in Auction A. When the price hits 50, bidders can choose between the auctions since they have the same standing high bid. The inexperienced bidder bids 51 in Auction A because bidding in Auction B would cost 50 + 10 (switching cost). Similarly, it is more profitable for the experienced bidder to continue bidding in Auction A.

However, when the price gets up to 55 in Auction A, the experienced bidder will be indifferent between switching to Auction B, whereas the inexperienced bidder will stay in
Auction A. An equilibrium is reached with the inexperienced bidder winning Auction A $55 and the experienced bidder winning Auction B at $50. The difference is $f_H = 5$, with the low-reserve auction achieving the higher ending price.

This result generalizes upward to the case where there are more than two auctions and more than two bidders. Thus, we should see that experienced bidders generally win auctions at lower prices than inexperienced bidders.

**Model 2: (Martingales)**

To answer Empirical Question #4, I investigate a version of Weber’s Martingale Theorem. In this model, I test the hypothesis that when several identical items are sold in sequence, the ending prices will form a martingale. That is, the expectation of the \( n + 1 \)st price, given the prior prices, is equal to the \( n \)th price. To formulate this problem, suppose there are \( k \) identical items for sale and \( N \) bidders. Each bidder has singleton demand. Now, let \( I_n \) denote the information that has become available after \( n \) items have been sold, and suppose that there is a symmetric, increasing equilibrium bid function \( \beta_{n+1}(\cdot | I_n) \) that applies to bidding for the \( n + 1 \)st item. At equilibrium, the order statistics imply that bidder with the highest type will win the first item; the bidder with the second highest type will win the second item, and so on. Let \( p_n \) denote the price of the \( n \)th item; \( t^{(1)}, \ldots, t^{(N)} \) denote the order statistics in decreasing order; and \( I_0 \) denote null information.

**Theorem:** At any equilibrium, the sequence of prices and information \((p_n, I_n)_{n=1}^k\) satisfies

\[
E[p_n | I_{n-1}] = E[v(t^{(k+1)}) | I_{n-1}].
\]  

(5-1)
If \( I_n \) is the sequence of past prices \( \{p_1, \ldots, p_n\} \), then \( (p_n, I_n)_{n=1}^k \) forms a martingale. (Milgrom 2004)

**Proof:** Assume bidder 1 has not won an item yet, and the first \( m-1 \) items have already been sold. Apply Myerson’s lemma to the game starting with the sale of item \( m \). The expected payment by bidder 1 given \( I_{m-1} \) must be identical to the Vickrey auction, since the decision outcome is the same:

\[
E \left[ \sum_{n=m}^{k} p_n 1_{[l_{n-1}(\omega) \leq t(n)]} \mid I_{m-1} \right] = E \left[ v(t^{(k+1)}) 1_{[l^{(n-1)}(\omega)< t^{(k+1)}]} \mid I_{m-1} \right]. \tag{5-2}
\]

Now, since \( m = k \), bidder 1 will win if

\[
t^1 = t^{(k)} \tag{5-3}
\]

and

\[
E[p_k \mid I_{k-1}] = E[v(t^{(k+1)}) \mid I_{k-1}]. \tag{5-4}
\]

By symmetry, the identity of the winning bidder for the \( n \)th item is independent of the price \( p_n \), so

\[
E[p_n 1_{[l_{n-1}(\omega)]} \mid I_{m-1}] = E[1_{[l_{n-1}(\omega)]} \mid I_{m-1}] E[p_n \mid I_{m-1}] = \frac{1}{N+1-m} E[p_n \mid I_{m-1}]. \tag{5-5}
\]

Similarly,

\[
E[v(t^{(k+1)}) 1_{[l^{(n-1)}(\omega)< t^{(k+1)}]} \mid I_{m-1}] = E[1_{[l^{(n-1)}(\omega)< t^{(k+1)}]} \mid I_{m-1}] E[v(t^{(k+1)}) \mid I_{m-1}] \]

\[
= \frac{k+1-m}{N+1-m} E[v(t^{(k+1)}) \mid I_{m-1}]. \tag{5-6}
\]

Therefore, these equations simplify to

\[
\frac{1}{N+1-m} \sum_{n=m}^{k} E[p_n \mid I_{m-1}] = \frac{k+1-m}{N+1-m} E[v(t^{(k+1)}) \mid I_{m-1}]. \tag{5-7}
\]

I conclude that
\[ E[p_m | I_{m-1}] = E[v(t^{(k+1)}) | I_{m-1}] \] for all \( m \leq n \leq k \), \hspace{1cm} (5-8)

because otherwise there would be some \( n^* \) that is the largest value of \( n \) for which the equality fails. A contradiction then arises by setting \( m = n^* \) in the last equation. Since

\[ E[p_m | I_{n^*}] = E[v(t^{(k+1)}) | I_{n^*}], \hspace{1cm} (5-9) \]

\( m \leq n^* \) implies that

\[ E[p_m | I_{m-1}] = E[E[p_n | I_{n^*}] | I_{m-1}] = E[E[v(t^{(k+1)}) | I_{n^*}] | I_{m-1}] = E[v(t^{(k+1)}) | I_{m-1}], \hspace{1cm} (5-10) \]

which is a contradiction.

By inverting the bid functions, the information \( I_n \) is \((t^{(1)}..t^{(n-1)})\), so by Myerson’s lemma

\[ E[p_m | I_{m-1}] = E[v(t^{(k+1)}) | I_{m-1}] = p_{m-1}. \hspace{1cm} (5-11) \]

Thus, given a set of previous auction prices, the expected ending price of the next auction is identical to the ending price of the previous auction. This is precisely the definition of a martingale, so the theorem holds.

6. Empirical Approach

My goal for this paper is to use empirical results to develop a theory of bidder behavior on eBay. This analysis will be performed with several different methods. First, ordinary least squares regressions will be used to verify basic properties of the data. These elementary analyses will confirm the results summarized in the literature review.

Then, for a more in-depth analysis of bidder behavior, I will apply McFadden’s choice model to my dataset. McFadden (1974) developed the theory to help make statistical inferences on a model of individual choice behavior from data obtained by sampling from a population of
individuals. This model, which uses conditional logit regressions to investigate interaction effects between variables, can be used to predict a bidder’s choices in a set of online auctions. It is worth mentioning that McFadden’s model can be applied to everything from finding markets for commodities to predicting voting behavior; for his work, he received the 2000 Nobel Prize in Economics.

In McFadden’s model, \( n \) individuals (i.e., the bidders) each make a single choice among \( J \) alternatives (i.e., all the auctions listed on one particular day). For the \( i \)th consumer, suppose that the utility of choice \( j \) is

\[
U_{ij} = z'_{ij} \beta + \varepsilon_{ij},
\]

where \( z'_{ij} \) contains other variables for the auction such as the current high bid and the amount of time until the auction closes, \( \beta \) is a vector of corresponding coefficients, and \( \varepsilon_{ij} \) denotes the random disturbance variable for each bidder choice.

The model assumes that the error terms in the model are independent and homoskedastic. If the consumer makes choice \( j \), we assume that \( U_{ij} \) is the maximum among the \( j \) utilities. Thus, the statistical model is driven by the probability that choice \( j \) is made:

\[
\Pr[U_{ij} > U_{ik}], \text{ for all other } k \neq j.
\]

McFadden’s model can be calculated by making an appropriate choice for the distribution of the \( \varepsilon_{ij} \) terms. The probit or logit models can be considered here. If probit is chosen, we must evaluate multiple integrals of the normal distribution. However, McFadden shows that it is much simpler to use a logit model.
First, we let $Y_i$ be a random variable that indicates the bidder’s choice. McFadden shows that if and only if the $J$ $\varepsilon_{ij}$ terms are independent and identically distributed with type I extreme value (Gumbel) distribution,

$$F(\varepsilon_{ij}) = \exp(-e^{-\varepsilon_{ij}}), \quad (6-3)$$

then

$$\Pr[Y_i = j] = \frac{e^{\beta z}}{\sum_{j=1}^{J} e^{\beta z}} , \quad (6-4)$$

which leads to what is called the conditional logit model (Greene 2003).

We can also make this model more specific by noting that some characteristics are specific to the individual as well as to the choices. In my dataset, these variables include a bidder’s preference of auctions with low reserves or a preference for auctions with a “Buy it Now!” designation. Let these variables be denoted $x_{ij}$. Other characteristics, such as a bidder’s feedback rating, are specific to the individual only. Call these variables $w_i$. Therefore, $z_{ij} = [x_{ij}, w_i]$. The critical difference between the two categories is that $x_{ij}$ varies across the choices for each bidder (e.g., time until the auction closes), while $w_i$ is the same for all choices faced by the bidder (e.g., the bidder’s experience level). In practice, this is accomplished by having a set of bid characteristics ($w_i$) duplicated $J$ times, with $x_{ij}$ characteristics merged in. Appendix E describes how to do this in Stata.

Now, if we incorporate $x_{ij}$ and $w_i$ into Equation 6-4, then we have

$$\Pr[Y_i = j] = \frac{e^{\beta x_{ij} + \alpha w_i}}{\sum_{j=1}^{J} e^{\beta x_{ij} + \alpha w_i}} = \frac{e^{\beta x_{ij}} e^{\alpha w_i}}{\sum_{j=1}^{J} e^{\beta x_{ij}} e^{\alpha w_i}}, \quad (6-5)$$
where $\beta$ is a vector of coefficients for $x$ and $\alpha$ is a vector of coefficients for $w$. This is the expanded form of McFadden’s choice model.

This thesis fills a gap in the literature by using this model to investigate interaction terms between bidders and auction characteristics. As the first paper to apply McFadden’s choice model to eBay auctions, it provides a new method for quantifying the relative effects of auction characteristics for bidder entry. By calculating odds ratios with interaction terms, I can discover tendencies among bidders, such as whether experienced bidders are more likely to bid in auctions run by experienced sellers or if experienced bidders are more likely to bid in auctions with a low reserve. The Results portion of this paper will discuss the model’s implementation and results in more detail.

7. Empirical Design

The dataset for this experiment consists of 3,135 auctions for invitations to the Google Gmail service, auctioned on eBay between 5/06/04 and 6/08/04, inclusive. These 3,135 auctions represent all invitations sold on eBay during the aforementioned time period, and thus provide interesting data for a wide variety of analyses beyond the scope of this paper. Auctions of several different types are included in the dataset—there are English-style ascending-price auctions, reserve-price auctions, fixed-price auctions, multiple-item listings, and auctions with a combination of these characteristics.

All the data were collected via eBay’s Completed Items Search, a tool that allows users to search for completed items ending in the past 15 days with a specified search string (in this case,
“Gmail”). Unfortunately, this search strategy resulted in a large number of false positives—it is not uncommon for eBay users to put oft-searched strings such as “Gmail” in auctions that are for unrelated or competing products to maximize exposure. Thus, auction pages were collected manually with Microsoft Internet Explorer and indexed by date. The bid history page was also collected for each auction, giving data on bidders, their bids, and their feedback ratings. The bid history pages were automatically collected using a Perl script. All auction and bid history pages were then parsed with a custom Java program to extract all necessary data. Finally, the data was imported into Microsoft Excel and Intercooled Stata 8.2 for regression analysis. Some of the code used for the data collection and analysis process is reproduced in Appendix C.

Although great care was taken to capture each auction to preserve the sequential nature of the data, several issues made this impossible. First, auctions with zero bids were deleted from the data. This was a reasonable course of action because amazingly, due to the hype surrounding Gmail’s launch, only a handful of auctions (roughly 0.2% of the data) closed with no winning bidder. Furthermore, there was generally an obvious reason for the absence of bids—nearly all of these auctions had an absurdly large minimum bid (i.e., $150 or more) that drove away all bidders.

Several auctions (and their corresponding bids) had to be manually deleted from the datasets when automatic parsing techniques failed. This occurred most frequently when the seller specified that bidder identities should be hidden. Since the datasets use the identities of the bidders, the only logical workaround was to ignore these auctions altogether.

Finally, auctions with fraudulent winners were ignored. A winner was defined as fraudulent if they received a negative feedback rating for the auction with some designation in the feedback comment that payment was never received. These auctions were ignored because
sellers have no control over the actions of bidders, and bidders who have no intention to pay frequently enter outlandish bids that create unnecessary noise in the dataset. On the other hand, since bidders see the identity and rating of all sellers, auctions with fraudulent sellers were not removed. Presumably, unreliable sellers have a low feedback rating, so bidder entry decisions would endogenize this risk.

After the dataset was cleaned up, summary statistics of the data were generated. Figure 1 shows auction statistics separated by day. The average ending price of the auctions started low (when few people knew about the service), rose steadily as the hype grew, and finally began to decrease as Google increased the supply of invitations. Not long after the data in this paper were collected, supply had increased so much that auctions usually ended for under $1.00 each.
<table>
<thead>
<tr>
<th>Auction End Date</th>
<th>Observations</th>
<th>Average End Price</th>
<th>Standard Deviation</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>05/06/04</td>
<td>114</td>
<td>$44.88</td>
<td>15.61</td>
<td>$26.00</td>
<td>$120.00</td>
</tr>
<tr>
<td>05/07/04</td>
<td>101</td>
<td>$40.30</td>
<td>13.69</td>
<td>$18.00</td>
<td>$105.50</td>
</tr>
<tr>
<td>05/08/04</td>
<td>79</td>
<td>$35.26</td>
<td>8.90</td>
<td>$17.50</td>
<td>$60.00</td>
</tr>
<tr>
<td>05/09/04</td>
<td>67</td>
<td>$32.82</td>
<td>7.34</td>
<td>$15.00</td>
<td>$61.00</td>
</tr>
<tr>
<td>05/10/04</td>
<td>115</td>
<td>$36.42</td>
<td>7.76</td>
<td>$21.50</td>
<td>$86.00</td>
</tr>
<tr>
<td>05/11/04</td>
<td>104</td>
<td>$37.72</td>
<td>8.91</td>
<td>$9.99</td>
<td>$81.00</td>
</tr>
<tr>
<td>05/12/04</td>
<td>145</td>
<td>$31.15</td>
<td>6.38</td>
<td>$13.99</td>
<td>$51.00</td>
</tr>
<tr>
<td>05/13/04</td>
<td>120</td>
<td>$38.97</td>
<td>15.49</td>
<td>$16.07</td>
<td>$150.00</td>
</tr>
<tr>
<td>05/14/04</td>
<td>60</td>
<td>$37.92</td>
<td>10.43</td>
<td>$23.73</td>
<td>$99.00</td>
</tr>
<tr>
<td>05/15/04</td>
<td>40</td>
<td>$41.51</td>
<td>7.70</td>
<td>$24.95</td>
<td>$71.00</td>
</tr>
<tr>
<td>05/16/04</td>
<td>35</td>
<td>$39.92</td>
<td>9.84</td>
<td>$12.50</td>
<td>$66.50</td>
</tr>
<tr>
<td>05/17/04</td>
<td>48</td>
<td>$37.38</td>
<td>15.31</td>
<td>$21.50</td>
<td>$125.00</td>
</tr>
<tr>
<td>05/18/04</td>
<td>70</td>
<td>$44.65</td>
<td>12.35</td>
<td>$24.99</td>
<td>$81.00</td>
</tr>
<tr>
<td>05/19/04</td>
<td>71</td>
<td>$50.32</td>
<td>12.00</td>
<td>$30.00</td>
<td>$102.50</td>
</tr>
<tr>
<td>05/20/04</td>
<td>140</td>
<td>$59.52</td>
<td>25.46</td>
<td>$35.00</td>
<td>$182.50</td>
</tr>
<tr>
<td>05/21/04</td>
<td>143</td>
<td>$70.49</td>
<td>23.28</td>
<td>$30.00</td>
<td>$199.90</td>
</tr>
<tr>
<td>05/22/04</td>
<td>111</td>
<td>$72.17</td>
<td>19.55</td>
<td>$0.99</td>
<td>$151.90</td>
</tr>
<tr>
<td>05/23/04</td>
<td>95</td>
<td>$56.46</td>
<td>13.20</td>
<td>$36.00</td>
<td>$103.50</td>
</tr>
<tr>
<td>05/24/04</td>
<td>124</td>
<td>$64.45</td>
<td>17.89</td>
<td>$35.00</td>
<td>$150.00</td>
</tr>
<tr>
<td>05/25/04</td>
<td>88</td>
<td>$70.92</td>
<td>15.24</td>
<td>$41.00</td>
<td>$111.90</td>
</tr>
<tr>
<td>05/26/04</td>
<td>114</td>
<td>$68.05</td>
<td>14.91</td>
<td>$40.00</td>
<td>$141.10</td>
</tr>
<tr>
<td>05/27/04</td>
<td>133</td>
<td>$71.64</td>
<td>17.17</td>
<td>$20.50</td>
<td>$150.00</td>
</tr>
<tr>
<td>05/28/04</td>
<td>98</td>
<td>$74.74</td>
<td>21.20</td>
<td>$46.00</td>
<td>$217.50</td>
</tr>
<tr>
<td>05/29/04</td>
<td>52</td>
<td>$69.03</td>
<td>12.30</td>
<td>$22.50</td>
<td>$90.00</td>
</tr>
<tr>
<td>05/30/04</td>
<td>46</td>
<td>$71.40</td>
<td>9.36</td>
<td>$50.00</td>
<td>$95.00</td>
</tr>
<tr>
<td>05/31/04</td>
<td>89</td>
<td>$74.25</td>
<td>20.74</td>
<td>$46.00</td>
<td>$152.50</td>
</tr>
<tr>
<td>06/01/04</td>
<td>47</td>
<td>$87.51</td>
<td>33.46</td>
<td>$24.99</td>
<td>$205.50</td>
</tr>
<tr>
<td>06/02/04</td>
<td>75</td>
<td>$73.89</td>
<td>14.60</td>
<td>$45.44</td>
<td>$119.90</td>
</tr>
<tr>
<td>06/03/04</td>
<td>66</td>
<td>$84.64</td>
<td>21.94</td>
<td>$44.00</td>
<td>$152.50</td>
</tr>
<tr>
<td>06/04/04</td>
<td>75</td>
<td>$75.26</td>
<td>17.50</td>
<td>$25.00</td>
<td>$142.50</td>
</tr>
<tr>
<td>06/05/04</td>
<td>39</td>
<td>$71.54</td>
<td>10.73</td>
<td>$59.99</td>
<td>$112.50</td>
</tr>
<tr>
<td>06/06/04</td>
<td>61</td>
<td>$79.29</td>
<td>20.37</td>
<td>$29.00</td>
<td>$146.80</td>
</tr>
<tr>
<td>06/07/04</td>
<td>75</td>
<td>$73.47</td>
<td>29.29</td>
<td>$39.99</td>
<td>$177.50</td>
</tr>
<tr>
<td>06/08/04</td>
<td>293</td>
<td>$41.13</td>
<td>17.86</td>
<td>$14.99</td>
<td>$119.90</td>
</tr>
</tbody>
</table>

Totals: 3133 $56.01 23.84 $0.99 $217.50

We also see that prices exhibited large price swings throughout the entire dataset. One fact not shown by the table is that nearly all the low-revenue auctions (auctions that ended with a closing price of $20 or below) were “Buy it Now!” auctions, implying that the ending price would have been much larger if the seller had not artificially capped her own revenue. Reasons for doing this could include the desire for a fast-closing auction to generate instant revenue, but more likely indicates a lack of sophistication and research by the seller. This is yet another
explanation for the well-documented phenomenon that experienced sellers generate higher returns on their auctions.

We can also separate the data by hour and sum the data across all days:

![Graph showing #Auctions vs. #Bidders](image)

**Figure 2:** Supply vs. demand on auctions. The $R^2$ value is approximately 0.98.

These preliminary statistics show that the number of auctions listed at any given time is directly proportional to the number of bidders at the time. This interesting phenomenon allows us a bit of leeway in controlling for the supply vs. demand aspects of price formation—although the number of buyers and sellers at any given time may play a small role in determine price, we also know that there is no time interval where supply far exceeds demand or vice versa. Thus, we can simplify our analysis by ignoring market thickness issues and controlling only for the possibility
that the *types* of auctions (defined by the auction characteristics) and bidders (e.g., experienced or inexperienced, as determined by feedback rating) are different at any given time.

The regressions needed to test the hypotheses of this paper require the effects of all auction parameters to be separated as thoroughly as possible. Having identical products with no shipping costs simplifies this task enormously, as idiosyncratic auction characteristics can be analyzed more precisely. The following dummy variables were created using the parsed auction data:

- **Experienced bidder dummy**
  - Examine the rating of the winning bidder. 1 denotes a bidder rating above the median user rating, and 0 otherwise. In this dataset, the median bidder rating was 17.

- **Experienced seller dummy**
  - Examine the rating of the auction’s seller. 1 denotes a bidder rating above the median user rating, and 0 otherwise. In this dataset, the median seller rating was 48.

- **Long description dummy**
  - 1 if description is longer than the median description length, 0 otherwise.

- **Weekend dummy**
  - 1 if auction ends on a weekend, 0 otherwise.

- **Long duration dummy**
  - 1 if the auction lasts for 4 or more days, 0 otherwise.

- **True Auction dummy**
  - 1 if the auction was set with a low reserve price (under $10.00), 0 otherwise.

- **8 hour dummies**
  - Dummies were created for each 3 hour interval of the day starting at 12:00AM.
• 34 date dummies
  o The items are not perfectly homogeneous—as more Gmail accounts are sold, the most
desirable usernames are taken, so each subsequent invitation becomes slightly less
valuable. There may also be snob effects that lower the value of invitations as the
program expands. Thus, separating auctions by their end date helps eliminate these time-
based price variations.

The bid history pages were also parsed to find a list of all bidders (not just the winners)
along with dummies for whether they are experienced and whether they were a bid leader or bid
follower (a user is defined as a bid leader if they are the first to bid on a given auction).

8. Methodology and Results

Using the variables described in the Empirical Design, I run an ordinary least squares
regression to examine Empirical Question #1.

\[ X = Bx_1 + Sx_2 + Nx_3 + Dx_4 + Lx_5 + Wx_6 + Tx_7 + \sum_{i=1}^{8} h_i x_{i,7} + \sum_{j=1}^{34} d_j x_{j,15} + \varepsilon \]  (8-1)

where

\( X \) = ending bid price

\( B \) = experienced bidder dummy

\( S \) = experienced seller dummy

\( N \) = “Buy it Now!” dummy

\( D \) = long description dummy

\( L \) = long auction duration dummy

\( W \) = weekend dummy

\( T \) = true auction (low reserve) dummy
Several predictions immediately follow from theory. Following the entry costs model, \( B \) should be negative, because it costs experienced bidders less to switch to a different auction. Thus, on average they will win auctions at a lower price.

\( S \) should be positive, as discussed extensively by many authors. See the literature review for several empirical examples.

\( N \) should be negative, because the “Buy it Now!” option on auctions effectively cuts off bidding. If the “Buy it Now!” price is below at least one bidder’s valuation, then the auction will end at a lower price than if the auction had been a regular ascending English auction, since one of the bidders will buy the item at the “Buy it Now!” price. If the “Buy it Now!” price is above everyone’s valuation, then the auction will not end in a sale and thus will not be included in the dataset. For this reason, eBay generally discourages sellers from using the “Buy it Now!” option unless bidders need an immediate sale (such as selling tickets for an upcoming sporting event).

\( D \) should be positive because a longer description generally indicates sincerity by the seller. If a seller is more trustworthy, their auction will generally attract more bidders and the auction price will be driven up. In general, descriptions also give more information to buyers and make them more confident about their purchase decision. However, for this dataset this effect will be minimal because even a nominal description will let buyers know exactly what they are getting.

\( L \) should be positive, according to several empirical studies discussed in the literature review.
$W$ receives little treatment in the theory and depends on bidder characteristics. If users are more likely to bid on auctions while at work, then $W$ will be negative. However, if users bid on auctions in their spare time on nights and weekends when they are more able to control their schedules, then $W$ will be positive.

$T$ may be positive or negative, as discussed earlier. Recall that McAfee, Quan, and Vincent proposed a theoretical model showing that auctions with high reserve prices generally end with a higher ending price. However, by changing the assumptions (as I have done in the entry costs model put forth in this paper), it is possible that low-reserve auctions will generate higher seller revenue.
Number of obs: 3164  
R-squared: 0.5393  
Adj R-squared: 0.5323  
Root MSE: 16.659

| Coeff. | Std. Err | t | P>|t| | [95% Conf. Interval] |
|--------|----------|---|------|------------------|
| **Dependent Variable:** endprice |
| expbidder | -1.618382 | 0.6648613 | -2.43 | 0.015 | -2.922 -0.314771 |
| expseller | 1.327953 | 0.7039011 | 1.89 | 0.059 | -0.0522 2.70811 |
| buyitnow | -4.824928 | 1.032494 | -4.67 | 0.000 | -6.8494 -2.800492 |
| longdescription | 2.870551 | 0.6487266 | 4.42 | 0.000 | 1.59858 4.142526 |
| longduration | 1.984444 | 1.130723 | 1.76 | 0.079 | -0.2326 4.201481 |
| binaryweekend | 70.55903 | 4.730134 | 14.92 | 0.000 | 61.2845 79.83352 |
| *trueauction* | -0.5496708 | 0.9609288 | -0.57 | 0.567 | -2.4338 1.334447 |

<table>
<thead>
<tr>
<th>Time Dummies</th>
</tr>
</thead>
<tbody>
<tr>
<td>hoursdum1</td>
</tr>
<tr>
<td>hoursdum2</td>
</tr>
<tr>
<td>hoursdum3</td>
</tr>
<tr>
<td>hoursdum4</td>
</tr>
<tr>
<td>hoursdum5</td>
</tr>
<tr>
<td>hoursdum6</td>
</tr>
<tr>
<td>hoursdum7</td>
</tr>
<tr>
<td>hoursdum8</td>
</tr>
<tr>
<td>Day1</td>
</tr>
<tr>
<td>Day2</td>
</tr>
<tr>
<td>Day3</td>
</tr>
<tr>
<td>Day4</td>
</tr>
<tr>
<td>Day5</td>
</tr>
<tr>
<td>Day6</td>
</tr>
<tr>
<td>Day7</td>
</tr>
<tr>
<td>Day8</td>
</tr>
<tr>
<td>Day9</td>
</tr>
<tr>
<td>Day10</td>
</tr>
<tr>
<td>Day11</td>
</tr>
<tr>
<td>Day12</td>
</tr>
<tr>
<td>Day13</td>
</tr>
<tr>
<td>Day14</td>
</tr>
<tr>
<td>Day15</td>
</tr>
<tr>
<td>Day16</td>
</tr>
<tr>
<td>Day17</td>
</tr>
<tr>
<td>Day18</td>
</tr>
<tr>
<td>Day19</td>
</tr>
<tr>
<td>Day20</td>
</tr>
<tr>
<td>Day21</td>
</tr>
<tr>
<td>Day22</td>
</tr>
<tr>
<td>Day23</td>
</tr>
<tr>
<td>Day24</td>
</tr>
<tr>
<td>Day25</td>
</tr>
<tr>
<td>Day26</td>
</tr>
<tr>
<td>Day27</td>
</tr>
<tr>
<td>Day28</td>
</tr>
<tr>
<td>Day29</td>
</tr>
<tr>
<td>Day30</td>
</tr>
<tr>
<td>Day31</td>
</tr>
<tr>
<td>Day32</td>
</tr>
<tr>
<td>Day33</td>
</tr>
<tr>
<td>Day34</td>
</tr>
<tr>
<td>Constant</td>
</tr>
</tbody>
</table>

Figure 3: Results for ordinary linear regression modeled by Equation 8-1.
Figure 3 shows the results of the regression of Equation 8-1. The $B, S, N, D,$ and $L$ coefficients match the results given by the models of this paper and other studies. $W$, somewhat surprisingly, is strongly positive; this implies that there is more competition between bidders during the weekends. $T$ is slightly negative but not statistically significant, lending credence to the idea that reserve prices matter only in the context of ex ante assumptions about bidders.

The data also show that auctions ending in the morning hours generate the highest ending price, while the most inexpensive products can be found in the evening. This phenomenon would normally point to the possibility of arbitrage. However, recall the earlier discussion of the secondary market for Gmail invitations: since it is difficult to resell invitations, arbitrage opportunities are minimal. Thus, this trend in prices can persist in equilibrium. One possible explanation for the tendency for morning auctions to end with a high price is that the supply of these auctions is much lower, as shown in Figure 2. Therefore, since the number of bidders is always much larger than the number of available auctions, bidders with high valuations will compete for a reduced set of items. They must also outbid any bidders who have previously entered a proxy bid. The result is that morning auctions will be bid up and will end with a higher ending price than any other time during the day, when the supply of auctions is much higher.

To answer Empirical Question #3, I run a chi-squares test to investigate the interaction between experience and bidding. A two-column table must be created for this computation. The first column is a dummy variable that takes the value of 1 if the bidder is the first bidder of an auction (referred to as a bidLeader). The second column, experienced, takes the value of 1 if the bidder has an above-median feedback rating.
This list must be carefully constructed because each bid history file may list several bids by one user (for example, bidder A may start off the bidding and rejoin later after being outbid by bidder B). Thus, I count only one bid per user per auction. Because I am investigating entry characteristics, I record only the first bid that each bidder makes during the auction.

Figure 4 shows the results of the tabulation.

<table>
<thead>
<tr>
<th></th>
<th>experienced</th>
</tr>
</thead>
<tbody>
<tr>
<td>bidLeader</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>8,240</td>
</tr>
<tr>
<td>1</td>
<td>1,206</td>
</tr>
<tr>
<td>Total</td>
<td>9,446</td>
</tr>
</tbody>
</table>

Figure 4: Output of chi-squares tabulation. experienced is the dummy for bidder experience (1 for experienced bidder, 0 otherwise), and bidLeader is a dummy variable indicating if the bidder is the first bidder or not.

The chi-squares analysis indicates that as predicted by the model, there is statistically significant evidence showing that auctions are more likely to have an experienced bidder as bid leader; i.e., experienced bidders tend to avoid competition with other bidders. The analysis also shows that inexperienced bidders are more likely to be bid followers.

To further test Empirical Questions #1-3, I apply McFadden’s choice model to the dataset. McFadden’s methods, which were described earlier in this paper, allow us to quantify the factors that lead to bidder entry. For example, if the coefficient on dummy variable A is twice as large as the coefficient on dummy variable B, this implies that the presence of A is twice as influential on a bidder’s entry decision as the presence of B. These coefficients can also be given as odds ratios. For example, an odds ratio of 1.10 on a dummy variable means that the presence of this characteristic increases the chance of a bid by 10%.

The model is formulated as follows: a set of bidders logs onto eBay at some particular time and examines all auctions ending in the next three days (eBay’s minimum auction duration).
Then, based on each auction’s specific parameters, the bidders make a choice on which auction to bid in. This set of assumptions, by construction, lends itself well to McFadden’s choice model. The dataset will be grouped by the bid ID and will have \((\# \text{bids} \times \# \text{auctions})\) observations. Thus, since 2,285 bids were observed in 347 auctions over this three day time span, a dataset of \(2,285 \cdot 347 = 792,895\) observations was generated.

Figure 3 showed that experienced bidders generally pay less than inexperienced bidders. To study possible reasons for this phenomenon, I run McFadden’s model with several interaction terms generated from the experienced bidder dummy:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>buyitnow</td>
<td>dummy for “Buy it Now!” auction</td>
</tr>
<tr>
<td>trueauction</td>
<td>dummy for low-reserve auction</td>
</tr>
<tr>
<td>longduration</td>
<td>dummy for auction with duration &gt;4 days</td>
</tr>
<tr>
<td>expseller</td>
<td>dummy for experienced seller</td>
</tr>
<tr>
<td>longdescription</td>
<td>dummy for auction with long item description</td>
</tr>
<tr>
<td>closetoend</td>
<td>dummy: 1 if this bid would have represented a moderately late bid in this auction (1-30 minutes left in the auction)</td>
</tr>
<tr>
<td>lastminute</td>
<td>dummy: 1 if this bid would have represented a very late bid in this auction (&lt;1 minute left)</td>
</tr>
<tr>
<td>bidfollower</td>
<td>dummy: 0 if this bid would have been the first bid in the auction</td>
</tr>
<tr>
<td>hoursdum1</td>
<td>dummy: auction ends between 12AM-6AM</td>
</tr>
<tr>
<td>hoursdum2</td>
<td>dummy: auction ends between 6AM-12PM</td>
</tr>
<tr>
<td>hoursdum3</td>
<td>dummy: auction ends between 12PM-6PM</td>
</tr>
<tr>
<td>expdum1</td>
<td>experienced bidder * hoursdum1</td>
</tr>
<tr>
<td>expdum2</td>
<td>experienced bidder * hoursdum2</td>
</tr>
<tr>
<td>expdum3</td>
<td>experienced bidder * hoursdum3</td>
</tr>
<tr>
<td>bidderseller</td>
<td>experienced bidder * experienced seller</td>
</tr>
<tr>
<td>biddertrueauction</td>
<td>experienced bidder * trueauction</td>
</tr>
<tr>
<td>bidderlongduration</td>
<td>experienced bidder * longduration</td>
</tr>
<tr>
<td>expbidfollower</td>
<td>experienced bidder * bidfollower</td>
</tr>
<tr>
<td>expclosertoend</td>
<td>experienced bidder * closetoend</td>
</tr>
<tr>
<td>explastminute</td>
<td>experienced bidder * lastminute</td>
</tr>
<tr>
<td>expbin</td>
<td>experienced bidder * buyitnow</td>
</tr>
</tbody>
</table>
Italicized variables are interaction terms. The precise definitions of the `closetoend`, `lastminute`, and `bidfollower` variables are not intuitively obvious and thus merit an explanatory example.

**Example:** Suppose we have the following auctions:

<table>
<thead>
<tr>
<th>Auction</th>
<th>End Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>6:50</td>
</tr>
<tr>
<td>B</td>
<td>7:00</td>
</tr>
</tbody>
</table>

We observe the following bids in the data:

<table>
<thead>
<tr>
<th>Global Bid ID</th>
<th>Auction</th>
<th>Bidder</th>
<th>Bid Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>1</td>
<td>4:59</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>2</td>
<td>6:25</td>
</tr>
<tr>
<td>3</td>
<td>B</td>
<td>4</td>
<td>5:00</td>
</tr>
<tr>
<td>4</td>
<td>B</td>
<td>1</td>
<td>6:40</td>
</tr>
</tbody>
</table>

The construction of McFadden’s choice model will set `closetoend` equal to 1 if a bid at the specified bid time would have been between one and thirty minutes of the ending time for the auction in question. Similarly, `bidfollower` will be equal to 1 if there were already bids in the auction in question at the specified time of bid. The resulting dataset is below.

<table>
<thead>
<tr>
<th>Global Bid ID</th>
<th>Auction</th>
<th>Bidder</th>
<th>Bid Time</th>
<th>End Time</th>
<th>BidFollower</th>
<th>ClosetoEnd</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>1</td>
<td>4:59</td>
<td>6:50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>1</td>
<td>4:59</td>
<td>7:00</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>2</td>
<td>6:25</td>
<td>6:50</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>2</td>
<td>6:25</td>
<td>7:00</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>4</td>
<td>5:00</td>
<td>6:50</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>B</td>
<td>4</td>
<td>5:00</td>
<td>7:00</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>1</td>
<td>6:40</td>
<td>6:50</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>B</td>
<td>1</td>
<td>6:40</td>
<td>7:00</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
The lastminute variable is generated in the same way as closetoend, except that the dummy variable is set to 1 if there is less than one minute left in the auction.

The output of the application of McFadden’s model is shown in Figure 5. The data are also given as odds ratios for easier interpretation.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>buyitnow</td>
<td>-1.536555</td>
<td>0.219938 **</td>
<td>0.2151209</td>
<td>0.047313</td>
</tr>
<tr>
<td>trueauction</td>
<td>0.296545</td>
<td>0.126032 *</td>
<td>1.345204</td>
<td>0.169539</td>
</tr>
<tr>
<td>longduration</td>
<td>-0.585096</td>
<td>0.099786 **</td>
<td>0.557022</td>
<td>0.055586</td>
</tr>
<tr>
<td>expseller</td>
<td>-0.089731</td>
<td>0.091581 **</td>
<td>0.9141769</td>
<td>0.083721</td>
</tr>
<tr>
<td>longdescription</td>
<td>0.115407</td>
<td>0.075762</td>
<td>1.12233</td>
<td>0.08503</td>
</tr>
<tr>
<td>closetoend</td>
<td>3.680206</td>
<td>0.077815 **</td>
<td>39.65458</td>
<td>3.085702</td>
</tr>
<tr>
<td>lastminute</td>
<td>-0.532934</td>
<td>0.11643 **</td>
<td>0.5868806</td>
<td>0.068331</td>
</tr>
<tr>
<td>bidfollower</td>
<td>1.360575</td>
<td>0.106578 **</td>
<td>3.898434</td>
<td>0.415489</td>
</tr>
<tr>
<td>hoursdum1</td>
<td>0.129596</td>
<td>0.158995</td>
<td>1.138368</td>
<td>0.180995</td>
</tr>
<tr>
<td>hoursdum2</td>
<td>0.048681</td>
<td>0.102113</td>
<td>1.049885</td>
<td>0.107207</td>
</tr>
<tr>
<td>hoursdum3</td>
<td>-0.034796</td>
<td>0.101147</td>
<td>0.9658025</td>
<td>0.097688</td>
</tr>
<tr>
<td>expdum1</td>
<td>-0.165063</td>
<td>0.220846</td>
<td>0.8478403</td>
<td>0.187242</td>
</tr>
<tr>
<td>expdum2</td>
<td>-0.184493</td>
<td>0.141881</td>
<td>0.8315257</td>
<td>0.117978</td>
</tr>
<tr>
<td>expdum3</td>
<td>-0.121926</td>
<td>0.139696</td>
<td>0.8852139</td>
<td>0.123661</td>
</tr>
<tr>
<td>explong</td>
<td>-0.040059</td>
<td>0.107473</td>
<td>0.9607329</td>
<td>0.103252</td>
</tr>
<tr>
<td>bidderseller</td>
<td>0.126757</td>
<td>0.128169</td>
<td>1.135141</td>
<td>0.14549</td>
</tr>
<tr>
<td>biddertrueauction</td>
<td>0.333186</td>
<td>0.194824 *</td>
<td>1.395407</td>
<td>0.271859</td>
</tr>
<tr>
<td>explongduration</td>
<td>0.339093</td>
<td>0.132266 **</td>
<td>1.403674</td>
<td>0.185658</td>
</tr>
<tr>
<td>expbidfollower</td>
<td>-0.218287</td>
<td>0.143083</td>
<td>0.8038945</td>
<td>0.115024</td>
</tr>
<tr>
<td>expclosetoend</td>
<td>-0.260828</td>
<td>0.113694 *</td>
<td>0.7704136</td>
<td>0.087592</td>
</tr>
<tr>
<td>explastminute</td>
<td>0.172666</td>
<td>0.159629</td>
<td>1.18847</td>
<td>0.189714</td>
</tr>
<tr>
<td>expbin</td>
<td>0.520607</td>
<td>0.30419</td>
<td>1.683049</td>
<td>0.511967</td>
</tr>
</tbody>
</table>

* = 5% significance
** = 1% significance

Figure 5: Results for the 3-day version of McFadden’s choice model (experienced vs. inexperienced bidders).

Bidders may not have such a long time horizon, however. To investigate whether the results hold in a shorter time interval, the identical model was also run with a 2-day time frame.
Conditional (fixed-effects) logistic regression

<table>
<thead>
<tr>
<th>Dependent Variable: choice</th>
</tr>
</thead>
<tbody>
<tr>
<td>--------</td>
</tr>
<tr>
<td>buyitnow</td>
</tr>
<tr>
<td>trueauction</td>
</tr>
<tr>
<td>longduration</td>
</tr>
<tr>
<td>expseller</td>
</tr>
<tr>
<td>longdescription</td>
</tr>
<tr>
<td>closetoend</td>
</tr>
<tr>
<td>lastminute</td>
</tr>
<tr>
<td>bidfollower</td>
</tr>
<tr>
<td>hoursdum1</td>
</tr>
<tr>
<td>hoursdum2</td>
</tr>
<tr>
<td>hoursdum3</td>
</tr>
<tr>
<td>expdum1</td>
</tr>
<tr>
<td>expdum2</td>
</tr>
<tr>
<td>expdum3</td>
</tr>
<tr>
<td>explong</td>
</tr>
<tr>
<td>bidderseller</td>
</tr>
<tr>
<td>biddertrueauction</td>
</tr>
<tr>
<td>explongduration</td>
</tr>
<tr>
<td>expbidfollower</td>
</tr>
<tr>
<td>expclosetoend</td>
</tr>
<tr>
<td>explastminute</td>
</tr>
<tr>
<td>expbin</td>
</tr>
</tbody>
</table>

* = 5% significance
** = 1% significance

Figure 6: Results for the 2-day version of McFadden’s choice model (experienced vs. inexperienced bidders).

Time horizons shorter than two days are unlikely to deliver significant results because structural problems in the analysis of bid data become magnified. As Athey and Haile (2004) and others have noted, a substantial obstacle to such analysis is that the number of bidders cannot be observed on eBay. All models from classical auction theory use the key assumption that the number of bidders is known: that is, either the number of potential bidders or the number of bidders who have valuations above the reserve price. Since eBay auctions take place over several days and bidders become aware of the auction at different times during the auction, bidders who
log on and discover that the price has already risen past their valuation will not bid. Thus, such bidders will not be represented in the data.

With longer time horizons, however, the problem is alleviated because it is much more likely that these bidders will find another auction with a standing high bid that is below their reservation price. Accordingly, these bidders will be much more likely to appear in the dataset, and as shown in Figures 5 and 6, the results become much more statistically significant.

The above analysis classified bidders into experienced and inexperienced categories. However, this is not the only way to distinguish bidders—in fact, any attribute that meaningfully segments two types of bidders can be used in a similar conditional logit analysis.

To verify the results of the experienced bidder analysis, we can look at specific behavioral traits that might be associated with these bidders. Since I hypothesized earlier that experienced bidders have lower search costs and thus are more likely to switch between different auctions, I performed another iteration of McFadden’s choice model by classifying bidders according to their bidding habits.

Bidding habits can be segmented into two categories: some bidders may choose to bid early on an auction and switch to a different auction if they are outbid. Others may choose a bid sniping strategy, bidding only in auctions that are about to end. I thus define a dummy variable, freqbidder, for bidders who have placed bids in an above-median number of auctions. I then generate interaction terms with the same variables as in Figures 5 and 6:

---

Several alternatives could have been pursued here. For example, a dynamic set of auction choices could have been calculated for each bidder consisting of a specified number of auctions ending before and after the recorded bid. Of course, this introduces its own problems as well; since we don’t know when a bidder begins searching for an item, it is impossible to calculate an accurate set of choices for each buyer.
An analogous set of conditional logit regressions were performed on the same datasets as the ones analyzed previously, yielding the results given in Figures 7 and 8.

Analogous set of conditional logit regressions were performed on the same datasets as the ones analyzed previously, yielding the results given in Figures 7 and 8.

| Conditional (fixed-effects) logistic regression | Number of obs = 792895 |
| Dependent Variable: choice | Prob > chi2 = 0 |
| | Pseudo R2 = 0.1553 |

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>buyitnow</td>
<td>-1.343613</td>
<td>0.15192 **</td>
<td>0.2609012</td>
<td>0.039636</td>
</tr>
<tr>
<td>trueauction</td>
<td>-0.035418</td>
<td>0.104484</td>
<td>0.9652019</td>
<td>0.100848</td>
</tr>
<tr>
<td>longduration</td>
<td>-0.374322</td>
<td>0.090577 **</td>
<td>0.6877556</td>
<td>0.062295</td>
</tr>
<tr>
<td>expseller</td>
<td>0.022423</td>
<td>0.082694</td>
<td>1.022676</td>
<td>0.084569</td>
</tr>
<tr>
<td>longdescription</td>
<td>0.193934</td>
<td>0.071043 **</td>
<td>1.214016</td>
<td>0.086247</td>
</tr>
<tr>
<td>closetoend</td>
<td>4.049793</td>
<td>0.074323 **</td>
<td>57.38557</td>
<td>4.265092</td>
</tr>
<tr>
<td>lastminute</td>
<td>0.274013</td>
<td>0.095998 **</td>
<td>1.315232</td>
<td>0.12626</td>
</tr>
<tr>
<td>bidfollower</td>
<td>1.246884</td>
<td>0.100217 **</td>
<td>3.479485</td>
<td>0.348703</td>
</tr>
<tr>
<td>hoursdum1</td>
<td>0.088423</td>
<td>0.14118</td>
<td>1.09245</td>
<td>0.154232</td>
</tr>
<tr>
<td>hoursdum2</td>
<td>-0.240281</td>
<td>0.096291 *</td>
<td>0.7864073</td>
<td>0.075724</td>
</tr>
<tr>
<td>hoursdum3</td>
<td>-0.224237</td>
<td>0.09529</td>
<td>0.7991257</td>
<td>0.075541</td>
</tr>
<tr>
<td>freqdum1</td>
<td>-0.239363</td>
<td>0.23082</td>
<td>0.7871291</td>
<td>0.181685</td>
</tr>
<tr>
<td>freqdum2</td>
<td>0.370603</td>
<td>0.14298 **</td>
<td>1.448607</td>
<td>0.207122</td>
</tr>
<tr>
<td>freqdum3</td>
<td>0.253257</td>
<td>0.1402</td>
<td>1.288215</td>
<td>0.180608</td>
</tr>
<tr>
<td>freqlong</td>
<td>-0.193993</td>
<td>0.108648</td>
<td>0.8236639</td>
<td>0.08949</td>
</tr>
<tr>
<td>freqseller</td>
<td>-0.155545</td>
<td>0.130728</td>
<td>0.8559483</td>
<td>0.111896</td>
</tr>
<tr>
<td>freqtrueauction</td>
<td>1.677725</td>
<td>0.223633 **</td>
<td>5.353361</td>
<td>1.197188</td>
</tr>
<tr>
<td>freqlongduration</td>
<td>-0.033247</td>
<td>0.130486</td>
<td>0.9672994</td>
<td>0.126219</td>
</tr>
<tr>
<td>freqbidfollower</td>
<td>0.015429</td>
<td>0.141432</td>
<td>1.015549</td>
<td>0.143631</td>
</tr>
<tr>
<td>freqclosetoend</td>
<td>-1.118737</td>
<td>0.117225 **</td>
<td>0.3266921</td>
<td>0.038297</td>
</tr>
<tr>
<td>freqlastminute</td>
<td>-2.063791</td>
<td>0.203872 **</td>
<td>0.1269717</td>
<td>0.025886</td>
</tr>
</tbody>
</table>

* = 5% significance
** = 1% significance

Figure 7: Results for the 3-day version of McFadden’s choice model (frequent vs. infrequent bidders).
### Conditional (fixed-effects) logistic regression

Number of obs = 364148  
Prob > chi2 = 0  
Pseudo R2 = 0.1324

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>buyitnow</td>
<td>-1.535251</td>
<td>0.188197</td>
<td>0.2154017</td>
<td>0.040538</td>
</tr>
<tr>
<td>trueauction</td>
<td>-0.026206</td>
<td>0.132478</td>
<td>0.9741348</td>
<td>0.129051</td>
</tr>
<tr>
<td>longduration</td>
<td>-0.346849</td>
<td>0.105615</td>
<td>0.7069124</td>
<td>0.07466</td>
</tr>
<tr>
<td>expseller</td>
<td>0.189914</td>
<td>0.100102</td>
<td>1.209145</td>
<td>0.07466</td>
</tr>
<tr>
<td>longdescription</td>
<td>0.089953</td>
<td>0.0943</td>
<td>1.094122</td>
<td>0.103176</td>
</tr>
<tr>
<td>closetoend</td>
<td>3.604835</td>
<td>0.09823</td>
<td>36.77561</td>
<td>3.612461</td>
</tr>
<tr>
<td>lastminute</td>
<td>0.004807</td>
<td>0.118353</td>
<td>1.004819</td>
<td>0.118923</td>
</tr>
<tr>
<td>bidfollower</td>
<td>0.947546</td>
<td>0.128873</td>
<td>2.579372</td>
<td>0.332412</td>
</tr>
<tr>
<td>hoursdum1</td>
<td>0.08407</td>
<td>0.177385</td>
<td>1.087705</td>
<td>0.192943</td>
</tr>
<tr>
<td>hoursdum2</td>
<td>-0.364709</td>
<td>0.121419</td>
<td>0.694396</td>
<td>0.084313</td>
</tr>
<tr>
<td>hoursdum3</td>
<td>-0.37797</td>
<td>0.115764</td>
<td>0.665251</td>
<td>0.079328</td>
</tr>
<tr>
<td>freqdum1</td>
<td>0.051907</td>
<td>0.2776</td>
<td>1.053278</td>
<td>0.29239</td>
</tr>
<tr>
<td>freqdum2</td>
<td>0.691562</td>
<td>0.180115</td>
<td>1.996832</td>
<td>0.359659</td>
</tr>
<tr>
<td>freqdum3</td>
<td>0.536266</td>
<td>0.170615</td>
<td>1.709612</td>
<td>0.291685</td>
</tr>
<tr>
<td>freqlong</td>
<td>-0.064697</td>
<td>0.138748</td>
<td>0.937351</td>
<td>0.130056</td>
</tr>
<tr>
<td>freqseller</td>
<td>-0.29243</td>
<td>0.152023</td>
<td>0.746447</td>
<td>0.113477</td>
</tr>
<tr>
<td>freqtrueauction</td>
<td>1.522873</td>
<td>0.247829</td>
<td>4.585379</td>
<td>1.136388</td>
</tr>
<tr>
<td>freqlongduration</td>
<td>-0.041089</td>
<td>0.147899</td>
<td>0.959743</td>
<td>0.141945</td>
</tr>
<tr>
<td>freqbidfollower</td>
<td>-0.19283</td>
<td>0.182971</td>
<td>0.819318</td>
<td>0.149912</td>
</tr>
<tr>
<td>freqclosetoend</td>
<td>-1.05636</td>
<td>0.147069</td>
<td>0.347719</td>
<td>0.051139</td>
</tr>
<tr>
<td>freqlastminute</td>
<td>-1.745924</td>
<td>0.217309</td>
<td>0.174483</td>
<td>0.037917</td>
</tr>
</tbody>
</table>

* = 5% significance  
** = 1% significance

Figure 8: Results for the 2-day version of McFadden’s choice model (frequent vs. infrequent bidders).

Now, I use the results from Figures 5-8 as evidence for a general theory of bidder behavior.

**A Theory of Bidder Behavior in Private-Values Auctions on eBay**

All eBay bidders, regardless of type, use essentially the same methods to find and bid in auctions: after logging on to eBay, they enter a search string, examine a few auctions, and finally place a bid. The particular auction a bidder enters can be predicted based on her feedback rating and the particular choice set of auctions she has at the time.
If each of these bids were guaranteed to be a winning bid, then modeling bidder entry would be relatively simple. However, with a large market of bidders, it is inevitable that most potential buyers get outbid sometime during the auction. When this happens, bidders face a choice: do they keep entering successively higher bids in the same auction, hoping to outbid all other bidders, or do they give up and find a completely new auction to enter instead?

Experienced bidders, by virtue of their increased comfort with the eBay bidding system, have lower search costs and are therefore more willing to switch between several different auctions in order to find the best deal. They tend to be more patient; thus, these bidders tend not to commit to a single auction when they might find a better price elsewhere even if it means that they will have to wait longer. On the other hand, it is difficult for inexperienced bidders to use complex bidding strategies. Instead, they tend to stick with one auction and keep bidding until they win. The empirical study shows that experienced bidders derive positive gains with the strategy of switching between auctions to avoid competition with other bidders as much as possible.

The preceding figures allow us to make a number of specific observations about bidder behavior in the case study of Gmail invitations (and other auction markets where bidders have private values and singleton demand). First, note that the results are similar for both time horizons regardless of the method used to separate bidders. Now, using the coefficients from Figures 5 and 6, we can describe experienced bidders with the following characteristics:

- more likely to bid in a “Buy it Now!” auction
- more likely to either bid early or bid in the last minute, not between
- more likely to bid in an auction with a long duration
- more likely to bid in an auction with a low reserve price
• more likely to be the first bidder in an auction

All of these characteristics are profitable strategies in these circumstances; they are all used with the idea of minimizing bidder competition. As we know, “Buy it Now!” auctions end with the first bidder. Therefore, if the “Buy it Now!” price is low enough, bidders can avoid competition completely by purchasing the item immediately at the designated price. Sellers who place a “Buy it Now!” price generally also specify a low minimum bid; the “Buy it Now!” option disappears when the first bid is placed at the minimum price level. Thus, the “Buy it Now!” price binds only when it is sufficiently below a bidder’s valuation such that the bidder feels confident she will not be able to find a better price by using a regular ascending bidding strategy.

Early bids by experienced users can deter other bidders from entering a bid in the auction. Recall the entry costs model earlier in this paper: if it is common knowledge that inexperienced bidders have a higher entry cost than experienced bidders, then an inexperienced bidder will be more easily deterred from bidding in an auction in which an experienced bidder has already bid, knowing that she will be at a disadvantage. Deterrence will work best in low-reserve auctions, which generally have a long duration.

If an experienced bidder identifies an auction in which the standing high bid is below her valuation with under a minute left in the auction, then it is profitable to use a bid sniping strategy. By waiting as long as possible before entering her valuation as a proxy bid, the bidder is guaranteed nonnegative payoffs—if she wins, it will be at a price less than or equal to her valuation; otherwise, the winning bid will be above her valuation and a zero payoff is preferable. The literature review discussed a paper by Roth and Ockenfels (2002), which showed that bid sniping is justified for experienced bidders in common-values auctions because of the need to protect their values. With sequential private-values auctions, however, there is no need for
bidders to protect their values. Bid sniping could even be detrimental: if several bidders simultaneously decide to snipe the same auction, the ending price will be driven up since there are plenty of other auctions to bid on instead, presumably with lower auction prices.

However, recall that Schindler (2003) provided other reasons to engage in bid sniping that apply to private-values auctions as well. Even Roth and Ockenfels concede that late bidding may be a viable strategy to reduce competition. Thus, some bid sniping may still be justified.

The data support these models. The coefficients for the expclosetoend, explastminute, and expbidfollower variables show that experienced bidders are most likely to bid either very early in the auction (i.e. be the first bidder of an auction), or very late (in the last minute before the auction closes). There is strong evidence that experienced bidders are much more likely to switch between different auctions, and the data support the hypothesis that the primary goal of these bidders is to avoid competition and price wars.

Figures 7 and 8 examine the behavior of bidders who tend to switch between auctions frequently. I call these users “frequent bidders.” These bidders can be characterized by the following attributes:

- more likely to bid between the hours of 6AM-6PM Pacific Time
- more likely to bid in a low-reserve auction
- less likely to engage in bid sniping

Note that although the regressions are performed in the same way, the definition of “frequent bidder” is not exactly analogous to the definition of an “experienced bidder” since the “frequent bidder” designation is a specific behavioral characteristic that represents a particular set of strategies. Instead, I use the analysis of frequent bidders as a logical extension of the experienced
bidder analysis. Since I earlier showed that experienced bidders are more patient and are more likely to switch to different auctions in order to find the best deal possible, we can look more specifically at these bidders in particular and examine characteristics of the most patient bidders. This sheds more light into how bidders can benefit from their patience.

The analysis of frequent bidders shows that these bidders choose to generally bid early in low-price auctions. Since most frequent bidders are experienced, this result agrees with the earlier analysis. To avoid competition with other bidders, they must bid in a large number of auctions since their early bids are much less likely to win. Because last-minute bids are the most likely to win and bidders have singleton demand, it is unsurprising that bidders who bid in many auctions do not engage in bid sniping.

To test Empirical Question #4, we must look for a martingale in the dataset. A subset of the data was chosen by looking for a sequential set of observations in which the following property holds:

Let \( n_1, \ldots, n_k \) denote auctions 1, 2, \( \ldots, k \) and let \( t_i \) denote the ending time of auction \( n_i \). Now,

\[
\forall i, (t_{i+1} - t_i) \leq 10 \text{ minutes.}
\]

This is enforced because we must have a short time period between auctions, by definition of a martingale. The result is a dataset of 120 sequential auctions ending between 6:31pm and 10:51pm on June 8, 2004. Then, all “Buy it Now!” auctions were removed because no incremental bidding occurred in these auctions. The final dataset is composed of 71 true ascending auctions, numbered 1 through 71.

Lagged variables were then generated using the auction number as the time variable, and an ordinary least squares regression was performed. As with the other regressions, other auction
characteristics such as bidder experience and seller experience are also included to partial out these effects. The results are shown in Figure 9.

We see moderately evidence strong evidence for a martingale in the dataset. At standard confidence levels, the values for lag2, lag3, lag4, and lag5 are essentially zero. The first lag variable, lag1, is statistically significant at the 0.05 level. This, however, can be explained by noting that since the auctions end in rapid succession, bidders often don’t have enough time to react to an auction’s ending price to use this information for the next closing auction. With some extra time, though, bidders can use this information in formulating the price of the next ending auction. Thus, it is not unreasonable to expect the zero means to start occurring only with the second lag variable.
9. Conclusion

The goal of this paper is to examine the effects of auction parameters on the outcome of online auctions. Although many authors have analyzed correlation between these parameters and ending prices of auctions, this paper takes a step further by exploring bidder choice models. To motivate the choice model, ordinary linear regressions were performed to verify the results of previous authors.

The results of these regressions provided many statistically significant results. Both buyer and seller reputation affected the ending price of an auction—experienced sellers were presumably more trustworthy and therefore commanded higher prices, while experienced bidders tended to find better bargains in their search. As expected, the “Buy it Now!,” long description, and long auction duration dummies matched predictions discussed previously.

The dummy variable for a low-reserve auction delivered statistically insignificant results. As mentioned in the models of this paper, this coefficient could be either positive or negative depending on the assumptions made about bidder behavior.

Though most of the results of my preliminary regressions mirrored the conclusions of other papers, my observation that experienced bidders pay less has not been adequately investigated by other authors. Thus, as the main contribution of this paper, I used an application of McFadden’s choice model to perform a complex study about bidder behavior. The goal of this analysis was to uncover some of the strategies experienced bidders use to win auctions at a lower cost. My implementation of McFadden’s model gave statistically significant evidence to support a general theory of experienced bidder behavior. I found that in a private-values, sequential auction framework where bidders have no incentive to hide their value, experienced bidders
systematically avoid price wars by bidding early in auctions and switching to different auctions when they are outbid. These bidders are most likely to bid early (to deter competition from other bidders) or during the very last minute of the auction (when other bidders have little time to react). This result was verified when I segregated bidders based on their bidding behavior (frequent bidders vs. infrequent bidders).

Future iterations of the McFadden’s choice model implementation may make the model more robust by calculating a dynamic set of choices for each bidder instead of assuming a static set of choices. In addition, different cutoff levels could be chosen for dummy variables representing bidder experience, seller experience, late bidding, etc.

My analysis of price dispersion resulted in evidence of a martingale; that is, given a set of \( n \) prior prices, the expected price of auction \( n + 1 \) is equal to the \( n \)th price. This result, predicted by Weber, has not been found in several noteworthy empirical analyses of offline auctions. However, my dataset supports the theory. This result could have been facilitated by the large number of nearly identical auctions available; bidders never needed to overpay in fear that they would miss out on the product.

This paper shows that bidders indeed gain from experience; thus, all bidders are not identical. Sellers can use this information to attract certain types of bidders to their auctions, depending on their desired objective. For example, if they simply want to maximize revenue, they can set up an auction with a long duration and long item description that ends on a weekend. If instead they want an experienced buyer to win the auction, setting a moderate “Buy it Now!” price may be the best method.

As auction theorist and Stanford professor Paul Milgrom is fond of saying, “The game is always bigger than you think.” Buyers, sellers, and auctioneers all have an incentive to
understand how auction parameters influence bidder behavior and auction prices because more information implies more power and flexibility to influence the outcome to their advantage. As I have shown, this extra information may allow participants to achieve a significantly better outcome for themselves than otherwise would have come about.
Appendix A: Sample Auction Page

<table>
<thead>
<tr>
<th>Gmail Account Invitation with Your Name in 1GB</th>
<th>Item number: 3679831416</th>
</tr>
</thead>
<tbody>
<tr>
<td>You are signed in</td>
<td></td>
</tr>
<tr>
<td>Bidding has ended for this item (jojoblance2001 is the winner)</td>
<td></td>
</tr>
</tbody>
</table>

Winning bid: US $61.01  
Ended: May 23, 04 07:21:00 PST  
Start time: May 20, 04 07:21:00 PST  
History: 10 bids  
Starting bid: US $9.95

Winner: jojoblance2001  

Go to larger picture

Sold by: Not a registered user  
Feedback Score: 32  
Positive Feedback: 100%  
Member since May 11, 04 in United States  
Read feedback comments  
Ask seller a question  
View seller's other items  

Item location: west United States  
Ships to: United States, South America, Asia, Europe, Canada

Shipping and payment details

**Description**

Get your own New Gmail account before it goes to public

I personally have my own account and I have to say, its pretty cool. I've been using it for a few weeks. You never have to search all e-mails in your account instead of going through all saved e-mails. Also, instead of a conversaion taking up several pages, gmail turns it into an easy to read thread. How nice! Get your's now!

Please include email address you'd like the invitation mailed to along with your paypal payment

I will send you the invitation for you to register your Gmail account with 1 gb. After one week, You may get invitations which you can give to your friends or sale it on ebay

The key features

- Gmail is an experiment in a new kind of webmail, built on the idea that you should never have to delete mail so you'll be able to find the message you want.
- Search, don't sort. Use Google search to find the exact message you want, no matter when it was sent or received.
- Don't throw anything away. 1GB megabytes of free storage so you'll never need to delete another message.
- Keep it all in context. Each message is grouped with all its replies and displayed as a conversation.
- No pop-up ads. No banners. You see only relevant text ads and links to related web pages of interest.

Policy and payment

- **Shipping is free**
- **There is no refund. All sale is final**

[Screen shot](http://gmail.google.com/gmail/help/receiv.html?from=vt&%20Your%20Item%20from%20Feature%20%3D%20Your%20Item%20from%20Feat)

Gmail User Reviews: [http://gmail.google.com/gmail/help/reviews.html](http://gmail.google.com/gmail/help/reviews.html)

Learn More: [http://www.google.com/gmail/help/learn_more.html](http://www.google.com/gmail/help/learn_more.html)

About Gmail: [http://gmail.google.com/gmail/help/about.html](http://gmail.google.com/gmail/help/about.html)
Appendix B: Corresponding Bid History Page

<table>
<thead>
<tr>
<th>User ID</th>
<th>Bid Amount</th>
<th>Date of bid</th>
</tr>
</thead>
<tbody>
<tr>
<td>joioclance/2001</td>
<td>US $61.01</td>
<td>May-23-04 07:18:21 PDT</td>
</tr>
<tr>
<td>joioclance/2001</td>
<td>US $61.01</td>
<td>May-23-04 07:16:02 PDT</td>
</tr>
<tr>
<td>joioclance/2001</td>
<td>US $61.01</td>
<td>May-23-04 07:12:55 PDT</td>
</tr>
<tr>
<td>demdec</td>
<td>US $60.01</td>
<td>May-23-04 07:17:35 PDT</td>
</tr>
<tr>
<td>fmaltziman</td>
<td>US $57.62</td>
<td>May-23-04 06:02:02 PDT</td>
</tr>
<tr>
<td>navinnawal</td>
<td>US $55.00</td>
<td>May-23-04 05:31:20 PDT</td>
</tr>
<tr>
<td>emed4u</td>
<td>US $50.00</td>
<td>May-23-04 05:41:34 PDT</td>
</tr>
<tr>
<td>ftrynnecs</td>
<td>US $50.00</td>
<td>May-23-04 05:19:06 PDT</td>
</tr>
<tr>
<td>ftrynnecs</td>
<td>US $47.00</td>
<td>May-23-04 05:18:55 PDT</td>
</tr>
<tr>
<td>navinnawal</td>
<td>US $45.00</td>
<td>May-22-04 23:00:52 PDT</td>
</tr>
<tr>
<td>navinnawal</td>
<td>US $33.00</td>
<td>May-22-04 22:40:05 PDT</td>
</tr>
<tr>
<td>nomad-dan</td>
<td>US $31.77</td>
<td>May-22-04 22:19:45 PDT</td>
</tr>
<tr>
<td>shote13</td>
<td>US $29.79</td>
<td>May-22-04 22:36:47 PDT</td>
</tr>
<tr>
<td>shote13</td>
<td>US $25.78</td>
<td>May-22-04 22:36:22 PDT</td>
</tr>
<tr>
<td>navinnawal</td>
<td>US $20.00</td>
<td>May-22-04 19:11:14 PDT</td>
</tr>
<tr>
<td>randyhallic</td>
<td>US $12.00</td>
<td>May-22-04 20:59:30 PDT</td>
</tr>
<tr>
<td>randyhallic</td>
<td>US $11.00</td>
<td>May-22-04 20:59:19 PDT</td>
</tr>
</tbody>
</table>

If you and another bidder placed the same bid amount, the earlier bid takes priority.
Appendix C: Auctioned Product

Invite a friend to join Gmail!
You have 1 invitation left

Thanks for using Gmail and helping us improve the service. We're ready to expand our test to a few more users, and because you've been a trusted early tester of Gmail, we're looking for your help. Please invite a few more people who you think would like Gmail and could help us make it even better.

Send Invite to:

First Name: __________________________
Last Name: __________________________
Email: ______________________________

Add a note to the invitation (optional):
I've been using Gmail and thought you might like to try it out. Here's an invitation to create an account.

preview invite text

Send Invite
Appendix D: Source Code

The source code in this section is a small subset of the code used to collect and prepare the data used for this paper. It is provided to give future researchers an idea for the tasks involved in a large-scale empirical study on an online marketplace such as eBay. All code is provided without support, maintenance, or warranty, and may be freely used or modified under the GNU General Public License <http://www.gnu.org/licenses/gpl.txt>.

```perl
#!/usr/bin/perl
use HTTP::Lite;
chdir "$/ARGV[0]" or die "cannot chdir to /$ARGV[0]: $!";

# bidhist.pl
# 09/27/04
# This script was run on a directory tree of auction files
# to collect the corresponding bid history pages.

foreach $i ($ARGV[2]..$ARGV[3]) {
    if (open (AUCTIONPAGE, "< $ARGV[0]$i.htm")) {
        @auctionpage = <AUCTIONPAGE>;
        $auctionpage = join '', @auctionpage;
        close AUCTIONPAGE;
        my @bidlink;
        $http = new HTTP::Lite;
        @bidlink = $auctionpage =~ m{(http://offer.ebay.com[^"]+)}gi;
        mkdir "C:\$ARGV[1]";
        chdir "C:\$ARGV[1]";
        open HISTORYFILE, "> b$i.htm" or die "Cannot open file: $!";
        if (@bidlink) {
            $bidlink[0] =~ s{&}{&}gi;
            $req = $http->request("$bidlink[0]"
                or die "Unable to get document: $!";
            print HISTORYFILE $http->body();
        }
        $http->reset();
        close HISTORYFILE;
    }
}
```
import java.util.*;
import java.io.*;
import java.util.regex.*;

public class BidAnalysis {

    private static class SingleBidder {
        String bidderName;
        int bidderRating;
        long itemNumber;
        int bidFollower;
        //File filename;
    }

    static int auctioncount = 0;

    public static void main(String[] args) throws IOException, InterruptedException {
        int count = 0;
        File dir = new File("../BidAnalysis/part1/");
        File[] dates = dir.listFiles();
        System.out.println("starting program");
        ArrayList auctionDB = new ArrayList();
        for (int i = 0; i < dates.length; i++) {
            if (dates[i].isDirectory()) {
                File[] auctions = dates[i].listFiles();
                for (int j = 0; j < auctions.length; j++) {
                    if (auctions[j].isFile()) {
                        //System.out.println(auctions[j].getAbsolutePath());
                        ProcessBidFile(auctionDB, auctions[j]);
                        count++;
                    }
                }
            }
            //if (count > 1000) break;
        }
        //sort by experience
        /*Collections.sort(auctionDB, new Comparator() { 
            public int compare(Object a, Object b) {
                SingleBidder first = (SingleBidder)a;
                SingleBidder second = (SingleBidder)b;
                if (first.bidderRating < second.bidderRating)
                    return -1;
                else if (second.bidderRating < first.bidderRating)
                    return 1;
                else return 0;
            }
        });*/
        //sort by auction number
        Collections.sort(auctionDB, new Comparator() { 
            public int compare(Object a, Object b) {
                SingleBidder first = (SingleBidder)a;
                SingleBidder second = (SingleBidder)b;
                if (first.itemNumber < second.itemNumber)
                    return -1;
                else if (second.itemNumber < first.itemNumber)
                    return 1;
                else return 0;
            }
        });
        //PrintTopBidders(auctionDB);
        System.out.println(auctionDB.size() + " bidders analyzed.");
        WriteToFile(auctionDB, "bidderDB.csv");
        System.out.println("count is " + count);
    }

    private static void PrintTopBidders(ArrayList auctionDB) {
        for (int i = 0; i < 100; i++) {
            System.out.println("bidder "+i+":" + auctionDB.get(i).bidderRating);
            if (i%5 == 0) System.out.println("------------------");
        }
    }
}

private static void ProcessBidFile(ArrayList auctionDB, File file) {
    try {
        BufferedReader reader = new BufferedReader(new FileReader(file));
        String line;
        while ((line = reader.readLine()) != null) {
            SingleBidder bidder = new SingleBidder();
            bidder.bidderName = line;
            bidder.bidderRating = Integer.parseInt(line);
            bidder.itemNumber = Long.parseLong(line);
            bidder.bidFollower = Integer.parseInt(line);
            auctionDB.add(bidder);
        }
        reader.close();
    } catch (IOException e) {
        System.out.println("Error reading file: "+file.getAbsolutePath());
    }
}
SingleBidder bidder = (SingleBidder) auctionDB.get(i);
System.out.println("bidder name: " + bidder.bidderName + "; bidder rating: " + bidder.bidderRating);
INUE "; filename: " + bidder.filename.getAbsolutePath());
}
}

private static void WriteToFile(ArrayList auctionDB, String filename) throws IOException, InterruptedException {
    BufferedWriter bidderDB = new BufferedWriter(new FileWriter(filename));
    System.out.println("size of database is " + auctionDB.size());
    bidderDB.write("username,dumExp,everyone,bidFollower,filename");
    bidderDB.newLine();
    for (int i = 0; i < auctionDB.size(); i++) {
        SingleBidder bidder = (SingleBidder) auctionDB.get(i);
        bidderDB.write(bidder.bidderName + ",");
        if (bidder.bidderRating > 20) bidderDB.write("1");  //cutoff
        else bidderDB.write("0");
        bidderDB.write(",1");
        bidderDB.write(",1" + bidder.bidFollower + "," + bidder.itemNumber);
        bidderDB.newLine();
    }
    bidderDB.close();
}

private static void ProcessBidFile(ArrayList auctionDB, File infile) throws FileNotFoundException, IOException {
    BufferedReader auctionFile = new BufferedReader(new FileReader(infile));
    String oneBidder = "<a \"ViewFeedback\&userid=[a-zA-Z0-9]*\\">\[0-9\]+";
    Pattern winningBidderRegex = Pattern.compile(oneBidder);
    Matcher matcher;
    boolean founditemnum = false;
    long itemnumber = 0;
    boolean firstbidder = true;
    while (true) {
        String nextLine = auctionFile.readLine();
        if (nextLine == null) break;
        nextLine = nextLine.trim();
        matcher = winningBidderRegex.matcher(nextLine);
        if (founditemnum == false) {
            if (nextLine.startsWith("Item number: <a href="*/Item\&item=")) {
                itemnumber = Long.parseLong(nextLine.substring(77,87));
                founditemnum = true;
                auctioncount++;
            }
        }
        if (matcher.find()) {
            SingleBidder singleAuction = new SingleBidder();
            String username = matcher.group().substring(7,
                matcher.group().lastIndexOf(">") + 2));
            if (ExistsAlready(auctionDB, username)) continue;
            int rating = Integer.parseInt(matcher.group().substring(matcher.group().lastIndexOf("\"\") + 2));
            singleAuction.bidderName = username;
            singleAuction.bidderRating = rating;
            //singleAuction.filename = infile;
            singleAuction.itemNumber = itemnumber;
            if (firstbidder) {
                singleAuction.bidFollower = 0;
                firstbidder = false;
            } else singleAuction.bidFollower = 1;
if (itemnumber == 0) System.out.println("couldn't find item num");
auctionDB.add(singleAuction);
}
auctionFile.close();
}
private static boolean ExistsAlready(ArrayList auctionDB, String username) {
  for (int i = 0; i < auctionDB.size(); i++) {
    SingleBidder oneUser = (SingleBidder)auctionDB.get(i);
    if (username.equals(oneUser.bidderName)) return true;
  }
  return false;
}

import java.util.*;
import java.io.*;
import java.util.regex.*;

public class AnalyzeAuctions {
  private static class OneAuction {
    double endPrice;
    double startPrice;
    int bidderRating;
    int sellerRating;
    File filename;
    boolean weekend;
    int endsInWeeHours;
    int intervalCode; //5/06-5/10: 1
    //5/11-5/15: 2
    //5/16-5/20: 3
    //5/21-5/25: 4
    //5/26-5/30: 5
    //5/31-6/04: 6
    //6/05-6/08: 7 //leave this one out so not linearly dependent
    int dayCode; //just put the day and use stata areg
    long auctionLength;
  }

  public static void main(String[] args) throws IOException, InterruptedException {
    int count = 0;
    File dir = new File(".");
    File[] dates = dir.listFiles();
    System.out.println("starting program");
    ArrayList auctionDB = new ArrayList();
    for (int i = 0; i < dates.length; i++) {
      File[] auctions = dates[i].listFiles();
      for (int j = 0; j < auctions.length; j++) {
        if (auctions[j].isFile()) {
          ProcessAuction(auctionDB, auctions[j]);
          count++;
        }
      }
    }
    //if (count > 1000) break;
    Collections.sort(auctionDB, new Comparator() {
      public int compare(Object a, Object b) {
        OneAuction first = (OneAuction)a;
        OneAuction second = (OneAuction)b;
        // compare the auction dates
        //...
OneAuction second = (OneAuction)b;
if (first.endPrice < second.endPrice)
    return 1;
else if (second.endPrice < first.endPrice)
    return -1;
else return 0;
}

PrintTopAuctions(auctionDB);
//CalculateStatistics(auctionDB);
WriteToFile(auctionDB, "auctionDB.csv");
System.out.println(count + " auctions analyzed.");

private static void WriteToFile(ArrayList auctionDB, String filename) throws IOException, InterruptedException {
    BufferedWriter bidderDB = new BufferedWriter(new FileWriter(filename));
    System.out.println("size of database is " + auctionDB.size());
    bidderDB.write("endprice, trueauction, expbidder, expseller, longdescription,
        binaryWeekend, endsInWeeHours, dayCode, dumIntervalTwo, dumIntervalThree, "+
        "dumIntervalFour, dumIntervalFive, dumIntervalSix, dumIntervalSeven,
        filename");
    bidderDB.newLine();
    for (int i = 0; i < auctionDB.size(); i++) {
        OneAuction auction = (OneAuction)auctionDB.get(i);
        int binaryWeekend = auction.weekend ? 1 : 0;
        int dumIntervalTwo = 0, dumIntervalThree = 0, dumIntervalFour = 0, dumIntervalFive = 0,
            dumIntervalSix = 0, dumIntervalSeven = 0;
        if (auction.intervalCode == 2) dumIntervalTwo = 1;
        else if (auction.intervalCode == 3) dumIntervalThree = 1;
        else if (auction.intervalCode == 4) dumIntervalFour = 1;
        else if (auction.intervalCode == 5) dumIntervalFive = 1;
        else if (auction.intervalCode == 6) dumIntervalSix = 1;
        else if (auction.intervalCode == 7) dumIntervalSeven = 1;
        int binaryLongDescription = (auction.auctionLength > 44000) ? 1 : 0;
        bidderDB.write(auction.endPrice + "," + auction.startPrice + "," + auction.bidderRating + "," + auction.sellerRating + "," + binaryLongDescription + "," +
            binaryWeekend + "," + auction.endsInWeeHours + "," + auction.dayCode + "," +
            dumIntervalTwo + "," + dumIntervalThree + "," + dumIntervalFour + "," + dumIntervalFive + "," + dumIntervalSix + "," + dumIntervalSeven + "," +
            auction.filename.getAbsolutePath());
        bidderDB.newLine();
    }
    bidderDB.close();
}

private static void ProcessAuction(ArrayList auctionDB, File infile) throws FileNotFoundException, IOException {
    //System.out.println(infile.toString());
    BufferedReader auctionFile = new BufferedReader(new FileReader(infile));
    OneAuction singleAuction = new OneAuction();
    String winningBidder = "ViewFeedback\&userid=[a-zA-Z0-9]*">[0-9]+";
    Pattern winningBidderRegex = Pattern.compile(winningBidder);
    String startPriceString = "[0-9]+,[0-9]+ starting bid";
    Pattern startPriceRegex = Pattern.compile(startPriceString);
    Matcher startPriceMatcher;
    Matcher userMatcher;
    boolean foundPrice = false, foundEndTime = false, foundStartPrice = false;
    int foundRating = 0;
    singleAuction.filename = infile;
    singleAuction.auctionLength = infile.length();
    while (true) {
        String nextLine = auctionFile.readLine();
        if (nextLine == null) {
            //System.out.println("reached end of file");
            auctionDB.add(singleAuction);
            break;
        }
        }
nextLine = nextLine.trim();
userMatcher = winningBidderRegex.matcher(nextLine);
startPriceMatcher = startPriceRegex.matcher(nextLine);

//regular auction
if (foundPrice == false && nextLine.startsWith("<TD width="100%"><FONT face=Arial size=2><B>US $")) {
    String price = nextLine.substring(48, 53);
    if (price.charAt(4) == '<')
        price = price.substring(0, 4);
    double priceNum = Double.parseDouble(price);
    //System.out.println("adding " + priceNum);
    singleAuction.endPrice = priceNum;
    foundPrice = true;
    if (priceNum > 150) System.out.println(infile.getAbsolutePath());
    //auctionDB.add(singleAuction);
    //break;
    //buy it now
} else if (foundPrice == false && nextLine.startsWith("<TD vAlign=top><FONT face=Arial size=2><B>US $")) {
    String price = nextLine.substring(46, 51);
    if (price.charAt(4) == '<')
        price = price.substring(0, 4);
    double priceNum = Double.parseDouble(price);
    //System.out.println("adding " + priceNum);
    singleAuction.endPrice = priceNum;
    foundPrice = true;
    //auctionDB.add(singleAuction);
    //break;
} else if (foundStartPrice == false && startPriceMatcher.find()) {
    double startPrice = Double.parseDouble(startPriceMatcher.group().substring(0, startPriceMatcher.group().length() - 13));
    //System.out.println("start price is " + startPrice);
    if (startPrice > 10)
        singleAuction.startPrice = 0;
    else singleAuction.startPrice = 1;
    foundStartPrice = true;
} else if (foundRating < 2 && userMatcher.find()) { //need to find buyer and seller
    int rating = Integer.parseInt(userMatcher.group().substring(userMatcher.group().lastIndexOf("">") + 2));
    //System.out.println(rating);
    if (foundRating == 0) { //buyer
        if (rating > 50) singleAuction.bidderRating = 1;
        else singleAuction.bidderRating = 0;
    } else { //seller
        if (rating > 50) {
            //System.out.println(foundRating + "seller rating is " + rating + ", coding 1 " + infile);
            singleAuction.sellerRating = 1;
        } else singleAuction.sellerRating = 0;
    } foundRating++;
} else if (foundEndTime == false && nextLine.startsWith("<TD><FONT face=Arial size=2>")) {
    int month, day, year;
    String endTime = nextLine.substring(28, nextLine.length() - 7);
    //System.out.println(endTime);
    int endHour = Integer.parseInt(endTime.substring(10, 12));
    if (endHour >= 2 && endHour <= 5)
        singleAuction.endsInWeeHours = 1;
    else singleAuction.endsInWeeHours = 0;
foundEndTime = true;
if (endTime.substring(0,3).equals("May"))
    month = 5;
else if (endTime.substring(0,3).equals("Jun"))
    month = 6;
else if (endTime.substring(0,3).equals("Apr"))
    month = 4;
else {
    System.err.println("Unexpected month found for end date");
    month = 0;
}
day = Integer.parseInt(endTime.substring(4,6));
year = 2004;
int endDayOfWeek = DayOfWeek(month-1, day, year);
if (endDayOfWeek == 6 || endDayOfWeek == 0)
    singleAuction.weekend = true;
else singleAuction.weekend = false;
if (month == 5) {
    singleAuction.dayCode = day - 5;
    if (day <= 10) singleAuction.intervalCode = 1;
    else if (day <= 15) singleAuction.intervalCode = 2;
    else if (day <= 20) singleAuction.intervalCode = 3;
    else if (day <= 25) singleAuction.intervalCode = 4;
    else if (day <= 30) singleAuction.intervalCode = 5;
    else singleAuction.intervalCode = 6;
} else if (month == 6) {
    singleAuction.dayCode = day + 26;
    if (day <= 4) singleAuction.intervalCode = 6;
    else singleAuction.intervalCode = 7;
} else {
    singleAuction.intervalCode = 0;
    singleAuction.dayCode = 0;
}
}
auctionFile.close();

private static int DayOfWeek( int month, int date, int year )
{
    int day;
    int mkeys[] = { 1, 4, 4, 0, 2, 5, 0, 3, 6, 1, 4, 6 };
    day = ( year - 1900 ) + ( year - 1900 ) / 4 + mkeys[month] + date - 1;
    /* The above counts the leap day even if it occurs later in the year */
    if(( year > 1900 ) && ( year % 4 == 0 ) && ( month < 2 ))
        day--;
    day %= 7;
    return day;
}
private static void PrintTopAuctions(ArrayList auctionDB) {
    for (int i = 0; i < 100; i++) {
        OneAuction auction = (OneAuction)auctionDB.get(i);
        System.out.println("ending price: " + auction.endPrice + "; bidder rating: " + auction.bidderRating + "; filename: " + auction.filename.getAbsolutePath());
    }
}
private static void CalculateStatistics (ArrayList auctionDB) {
    double max = 0, min = 1000;
    double sum = 0;
    double price;
    /*
     * for (int i = 0; i < auctionDB.size(); i++) {
     *     price = ((OneAuction)auctionDB.get(i)).endPrice;
     *     if (price > max) max = price;
     *     if (price < min) min = price;
     */
    sum += price;
}
    double avg = sum / auctionDB.size();

    System.out.println("max: " + max);
    System.out.println("min: " + min);
    System.out.println("avg: " + avg);
*/

    long length;
    int low = 0, high = 0;
    for (int i = 0; i < auctionDB.size(); i++) {
        length = ((OneAuction)auctionDB.get(i)).auctionLength;
        sum += length;
        if (length < 44472) low++;
        else high++;
    }
    double avg = sum / auctionDB.size();
    System.out.println("avg length: " + avg);
    System.out.println("below avg: " + low + ", above avg: " + high);
}
Appendix E: Stata Code

The code in this section was used to create the datasets and implement McFadden’s choice model. The source documents were bidderDB.csv, a worksheet with bidder-only variables (denoted $w_i$ in the theoretical portion of this paper), and auctionDB.dta, a worksheet for the other variables (denoted $x_{ij}$).

```stata
. insheet using bidderDB.csv
(10 vars, 793589 obs)
. sort auctionnumber
. merge auctionnumber using auctionDB-3day.dta
. gen bidfollower = (firstbidday < bidday | (firstbidday == bidday & firstbidminspastmidnight < bidminsaptmidnight > ))
. gen hoursdum1 = (hourscode==1 | hourscode==2)
. gen hoursdum2 = (hourscode==3 | hourscode==4)
. gen hoursdum3 = (hourscode==5 | hourscode==6)
. gen expdum1 = experienced * hoursdum1
. gen expdum2 = experienced * hoursdum2
. gen expdum3 = experienced * hoursdum3
. gen explong = experienced * longdescription
. gen bidderseller = experienced * expseller
. gen biddertrueauction = experienced * trueauction
. gen expbidfollower = experienced * bidfollower
. gen closetoend = (bidday == daycode & (minspastmidnight - bidminspastmidnight < 30) & (minspastmidnight - bidminspastmidnight > 1))
. gen lastminute = (bidday == daycode & (minspastmidnight - bidminspastmidnight < 1))
. gen expclosetoend = experienced * closetoend
. gen explastminute = experienced * lastminute
. gen expbin = experienced * buyitnow
. gen explongduration = experienced * longduration
. clogit choice buyitnow trueauction longduration expseller longdescription closetoend lastminute bidfollower hoursdum1 hoursdum2 hoursdum3 expdum1 expdum2 expdum3 explong bidderseller biddertrueauction explongduration expbidfollower expclosetoend explastminute expbin, group(id) or
(results in paper)
. clogit choice buyitnow trueauction longduration expseller longdescription closetoend lastminute bidfollower hoursdum1 hoursdum2 hoursdum3 expdum1 expdum2 expdum3 explong bidderseller biddertrueauction explongduration expbidfollower expclosetoend explastminute expbin, group(id) (results in paper)
. by username: egen bidfreq=count(experienced)
. gen bidfreq2 = bidfreq / 236
. gen freqbidder = (bidfreq2 > 10)
. gen freqlong = freqbidder * longdescription
. gen freqseller = freqbidder * expseller
. gen freqtrueauction = freqbidder * trueauction
. gen freqbidfollower = freqbidder * bidfollower
. gen freqclosetoend = freqbidder * closetoend
. gen freqbin = freqbidder * buyitnow
. gen freqlongduration = freqbidder * longduration
. gen freqlastminute = freqbidder * lastminute
. gen freqdum1 = freqbidder * hoursdum1
. gen freqdum2 = freqbidder * hoursdum2
. gen freqdum3 = freqbidder * hoursdum3
. clogit choice buyitnow trueauction longduration expseller longdescription closetoend lastminute bidfollower hoursdum1 hoursdum2 hoursdum3 freqdum1 freqdum2 freqdum3 freqlong freqseller freqtrueauction freqlongduration freqbidfollower freqclosetoend freqlastminute, group(id) or
. clogit choice buyitnow trueauction longduration expseller longdescription closetoend lastminute bidfollower hoursdum1 hoursdum2 hoursdum3 freqdum1 freqdum2 freqdum3 freqlong freqseller freqtrueauction freqlongduration freqbidfollower freqclosetoend freqlastminute, group(id)
References


