Consumer Surplus in Online Auctions

Ravi Bapna[+], Wolfgang Jank, Galit Shmueli[*]

July 13, 2006

ABSTRACT

Despite the growing research interest in Internet auctions, particularly those on eBay, little is known about quantifiable consumer surplus levels in such mechanisms. Using an ongoing novel field experiment that involves real bidders participating in real auctions, and voting with real dollars, we collect and examine a unique dataset to empirically quantify and understand determinants of consumer surplus in eBay auctions. The estimation procedure for private value auctions relies mainly on knowing the highest bid, which is not disclosed by eBay, but is available to us from our experiment. For common value auctions, where bidders bid strategically to avoid winner's curse, we develop an estimation procedure that infers the bidders' signals from their bids, and subsequently infers the item's common value and resulting surplus, from the signals. Our analysis, based on a sample of 4514 eBay auctions, indicates that the median surplus level per eBay auction is $3.61, which roughly translates to $1.5 billion in accrued consumer surplus for the year 2003 alone. On a percentage basis, consumers are capturing at least 18.3% of the total surplus generated in eBay auctions. We find that consumer surplus is significantly different across currencies and item categories, negatively influenced by seller experience, auction duration and competition, and positively influenced by bidder experience, bidder aggressiveness and item price.

Keywords: eBay, sniping, highest bid, Weibull distribution

I. INTRODUCTION

Classical microeconomic theory uses the notion of consumer surplus as the welfare measure that quantifies benefits to a consumer from an exchange. Alfred Marshall (1936) defined consumer surplus as "the excess of the price which he (a consumer) would be willing to pay rather than go without the thing, over that which he actually does pay…" It is also traditional to visualize consumer surplus as the roughly triangular area lying under a downward sloping demand curve and above the rectangle that represents actual money expenditure. Yet, despite its established theoretical standing, empirical studies of consumer surplus levels are not widely observed in the literature. In this paper we report on consumer surplus levels in eBay auctions, a vastly popular Internet based electronic market. eBay's popularity is evident in the reported $23.8 billion in gross merchandise sales for the year 2003, up from $14.9 billion in 2002. Quantifying the consumer surplus in eBay is one part of understanding its overall benefit to the economy.

To the best of our knowledge, this is the first attempt in IS research that deploys an IS artifact (in this case an eBay sniping agent developed by us called cniper.com) publicly and derives hitherto unobservable insights from data generated by the usage of the artifact. Our analysis is based on novel data from a field experiment that allows bidders to use our web based tool, cniper.com, to snipe eBay auctions. In using the tool, the bidders who end up winning the auction reveal to us, the otherwise unobserved, highest winning bid on eBay. This information provides a reliable way to estimate consumer surplus in both private and common value settings.

Internet based markets are now mainstream artifacts of today's economy. Yet, very little has been generally said about the quantifiable benefits such markets provide to consumers. While a variety of studies [Austan D. Goolsbee and Amil K. Petrin (2001), Aviv Nevo (2001)] have measured consumer surplus in traditional markets, such measures are scarce in electronic markets. There are two notable exceptions. Erik Brynjolfsson, Yu Hu and Michael D. Smith. (2003) in an interesting analysis, demonstrate how new product introduction in electronic markets can lead to significant consumer welfare gains. There is also an emerging stream of work, exemplified in Anindya Ghose, Michael D. Smith, and Rahul Telang (2004), which is looking closely at the welfare implications (including consumer welfare)

1

of secondary markets for used books. Both studies creatively devise econometric estimation procedures, based on Hausman (1981), to measure consumer welfare levels where they are not *directly observable* in posted price markets. Typically, in such markets, willingness to pay has to be inferred indirectly through surveys, contingent valuation techniques and price changing experiments such as promotions and discounts. Surveys of willingness to pay have credibility issues and have lead to a stream of research dealing with contingent valuation precision and bias reduction [see Peter A. Diamond and Jerry A. Hausman (1994)]. In this context, it is our belief that more needs to be said and done in a wider context about consumer surplus.

In this paper we demonstrate the suitability of using direct mechanisms [Roger Myerson (1981)], such as auctions, to quantify and understand determinants of consumer surplus. Auction theory is built upon the fact that a consumer with a valuation for an item, strategizes and formulates a bid, so as to maximize her surplus [R. Preston McAfee and John McMillan (1987)]. While auctions have been around for centuries, only in recent times, in conjunction with advancements in Internet technologies, have they reached the scope and scale to be considered mainstream for consumers. Lucking-Reiley (2000) provides an overview of what is being auctioned by whom and under what mechanism rules on the Internet. The growing popularity of Internet auctions has been accompanied with equal fervor amongst researchers revisiting auction theory, and finding new and creative uses for the vast amount of bidding data available. These range from classifying bidding strategies in multi-unit Yankee auctions [Ravi Bapna et al. (2004)], to determining optimal bid increments [Bapna et al. (2003)], to visualizing online auction data [Galit Shmueli and Wolfgang Jank (2004)] to studying the price dynamics in online auctions [Wolfgang Jank and Galit Shmueli (2003)]. For an excellent review of research on eBay in particular see Patrick Bajari and Ali Hortascu (2004).

Despite all this attention to eBay auctions, there are no published studies yet of consumer surplus generated in eBay. We are aware of two other groups of researchers currently involved in estimating consumer surplus levels generated in eBay. Tugba Giray, Kevin Hasker and Robin Sickles (2005) study data from auctions of computer monitors on eBay and estimate bidding functions by maximum likelihood

using five (Log-normal, Gamma, Weibull, Pareto, and Logistic) different assumptions about the underlying distribution of independent private values. Unjy Song (2004) develops a semi-parametric approach, and applies it to auctions of university yearbooks. A key feature of this approach is that it relies on the second and third highest bids observed on eBay to estimate the highest bid. In contrast, our study is designed to rely on the revealed willingness-to-pay of real bidders voting with their dollars in ebay auctions. Both [Giray, Hasker and Sickles (2005) and Song (2004)] studies assume a private value setting and provide valuable information for surplus levels in specific item categories of eBay. We view our work as providing an alternative broader view of consumer surplus generated in eBay. It should be noted, that eBay is a generalized auction house, carrying auctions in 30 major categories ranging from antiques to video games. This motivates us to develop estimation procedures for both private and common value eBay auctions. Our research objective is to present consumer surplus estimation procedures, as well as data, that reflect the wide variety of auctions on eBay.

We expect to contribute to this stream of research by addressing the following research questions:

i) How can consumer surplus be estimated in private and common value eBay auctions?

ii) What is the level and distribution of consumer surplus in eBay auctions? and

iii) How do auction, item, and market characteristics affect consumer surplus in eBay auctions? To study this we need to explicitly take into account the characteristics of an auction. The characteristics studied are bidder and seller experience, item category, currency, and seller mechanism design choices, such as opening bid, usage of a hidden reserve and auction duration.

We are motivated to quantify the level and distribution of consumer surplus not just to provide confidence intervals on the actual dollar levels of consumer surplus in eBay auctions, but also to establish robust benchmarks that can be used to measure the impact of future policy changes on consumer welfare.

These policy changes could range from increased eBay fees[1], to say, a new bid increment policy[2]. In addition, the value of characterizing an empirical distribution of consumer surplus lies in guiding current and future analytical work, that models the dynamics of eBay auctions, towards making valid distributional assumptions about bidder's valuations. As is evident from the work of Giray, Hasker and Sickles (2005), this is still an open research question.

The motivation behind understanding effects of auction/item characteristics on surplus levels is multi-fold. To the extent sellers view consumer surplus as money left on the table by them, do more experienced sellers leave less money on the table? In the same vein, do more experienced bidders extract higher surplus levels? All else being equal, do surplus levels vary systematically across item categories or across continents? How do seller mechanism design choices such as the opening bid, the usage of a hidden reserve and the choice of an auction's duration impact the winning bidder's surplus?

It is also worth discussing the usefulness and limitations of consumer surplus in informing us about the value of eBay as an exchange mechanism. Clearly, the overall value of eBay as an exchange is equal to the buyer's willingness to pay minus the seller's willingness to sell. We, and every other eBay study we are aware of, are limited by not having access to seller's willingness to sell. However, to the extent that sellers selling on eBay are behaving rationally, it is reasonable to assume that they have factored in their best outside option before deciding to sell on eBay. In equilibrium, this should be reflected in their reservation price, as well as in eBay's ability to charge for their services. If this is the case, consumer surplus can be added to the auction price to yield an upper bound on social surplus, or alternatively, a lower bound on the percentage of total surplus that is accrued to consumers. This stems from the fact that the auction price contains the other pieces of information needed to estimate social surplus. It certainly contains eBay's gains, as those map directly to price, albeit in two parts (fixed fee and commission). The other components of the price can be broken down into seller's surplus and seller's

---

[1] eBay attempted to significantly increase fees in early 2005, but had to retract due to widespread seller furor (see Broersma (2005), "Amid Customer Backlash, eBay Reduces Some Fees," available at http://www.eweek.com/article2/0,1895,1761416,00.asp)

[2] For instance, we were contacted by William Hsu of Ebay Product Marketing in June 2004 for advice on how to set dynamic bid increments.

reservation price. While auction price is readily posted for all to see on eBay, consumer surplus is not directly evident, and hence the focus of this research. We estimate that consumers are capturing at least 18.3%[3] of the total surplus generated in eBay auctions

I. A. Field Experiment to Determine Highest Winning Bid on eBay

There have been several studies that have described, in depth, eBay's second-price ascending proxy-bid auction mechanism. A key feature is that at the termination of the auction, the highest bidder wins and pays a price equal to the second highest bid plus one bid increment. The exceptions to this are when i) the two highest bids are equal, wherein the earlier bidder is awarded the item at a price equal to her bid; ii) the two highest bids are less than an increment apart, wherein the higher bidder is awarded the item at a price equal to her bid; iii) if the reserve price is higher than the second highest bid, wherein the higher bidder is awarded the item at the reserve price and if the Buy It Now is accepted. Another established feature, resulting primarily from eBay's hard closing time, is that last minute bidding or sniping is widely prevalent. Al Roth and Axel Ockenfels (2002) provide interesting theoretical and empirical insights into sniping behavior on eBay. They observe that in 240 antique auctions on eBay, 89 had bids in the last minute and 29 in the last 10 seconds.[4] Similar findings have been reported by Bajari and Hortaçsu (2003), Shmueli, Russo and Jank (2004) and Schindler (2003). Explanations for late bidding range from tacit collusion against sellers to the presence of naïve bidders who don't understand proxy bidding, to common value components in the items being auctioned. For the purpose of this study, we make use of the fact that sniping is prevalent.

eBay posts almost the complete bid history after the auction closes, *with the exception being the value of the highest bid.* For instance, consider the bid history of the following auction for a Nokia 6610 GSM cell phone. Conspicuous in its absence is the exact amount bid by the winner 'kanchenjunga[5].'

---

[3] Based on median surplus and price levels.
[4] Our own conversations with a senior eBay executive reveal that 80 percent of all bids arrive in the last hour of the auction.
[5] The first author's eBay id.

Since the winner's bid is not disclosed by eBay, there is no direct measure of the revealed willingness to pay of the winning bidder from the auction that is publicly available. To overcome this limitation, we design an ongoing field experiment that allows real-world bidders to use Cniper.com, our web based bidding agent, to snipe eBay auctions.
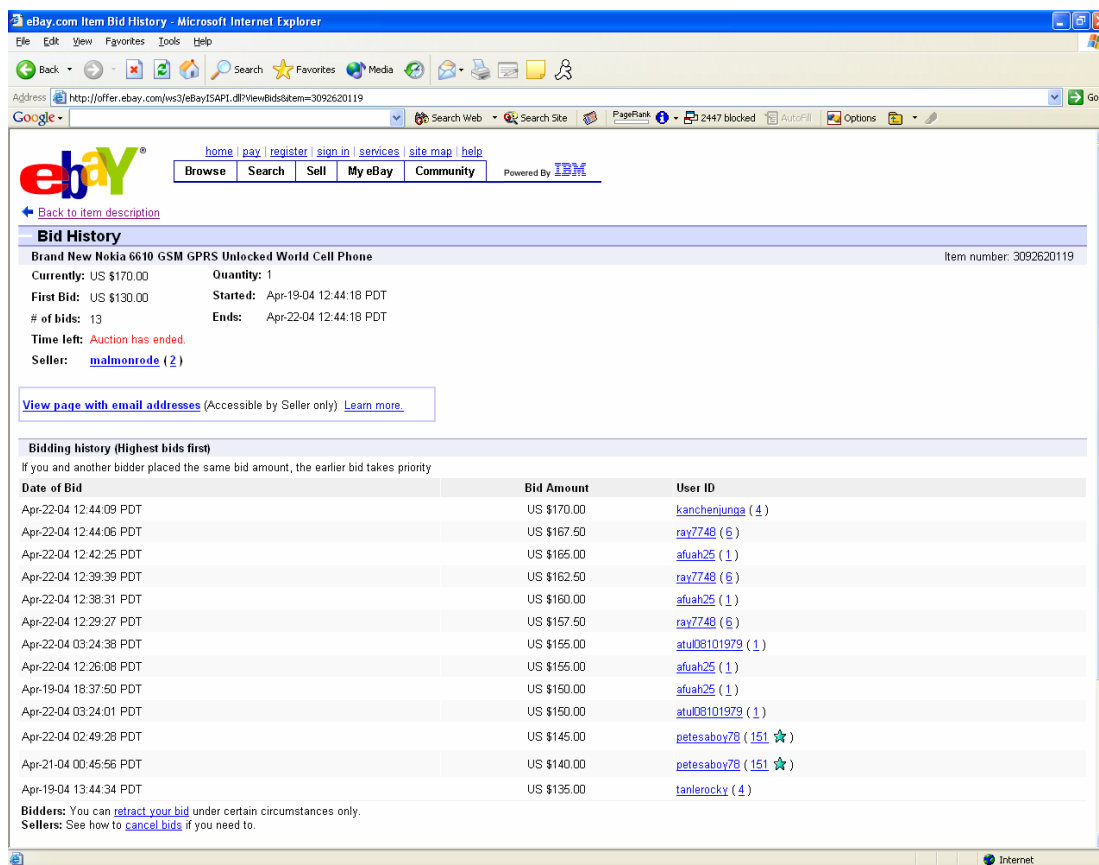
Figure 1 – Almost complete bid history, exception being the highest bid

Internet-based field experiments that deal with real bidders in real markets provide a contrast to the controlled environment of laboratory experiments with student subjects. This is evident in the work of David Lucking-Reiley[6] (1999) and John A. List and Lucking-Reiley (2002). They show how age old questions such as revenue equivalence and the importance of decisions costs respectively, can be practically examined using field experiments with real bidders and without any theoretical assumptions

---

[6] Now David Reiley.

6

that would be enforced in the laboratory. The current study is designed in the spirit of a field-experiment, with real bidders participating in real auctions, and voting with real dollars.

Bidders using Cniper.com to bid on their behalf reveal their willingness to pay to the agent. For auctions where our agent wins the auction, we can measure surplus in private and common value settings, if we have an appropriate model that subtracts from the item's value the actual price paid.

I. B. Model for Estimating Consumer Surplus

We consider the general setup described in Paul Klemperer (1999, page 58) and Vijay Krishna (2002, Section 6.1) where bidder i has a signal denoted by x_i and a valuation v_i. Following the notation of Klemperer, we have

$$v_i = \alpha x_i + \beta \sum_{j \neq i} x_j , \qquad (1)$$

where α and β, ($\alpha \geq \beta$), are weighting components that indicate the degree of private/common value component. Using the lens of William Vickrey's (1961) stylized model, eBay's mechanism is a hybrid between an open ascending English auction and a sealed bid second price auction. For such hybrid mechanisms multiple equilibria are likely to exist and are being currently explored [Bajari and Hortascu (2003), Hasker et al. (2005)].

I. B.I Case – Independent Private Values

The case $\beta=0$ gives the private value model. In that model, bidder's are concerned only about their own signals and not signals of others. This indicates that bidder types can be represented by their valuations, thus $x_i = v_i$. Let $v_{(1)}$ denote the highest valuation amongst the pool of bidders in an auction. Let p denote the auction closing price as observed on eBay. In the sniping stages of the eBay auction, under the private values setting, the absence of any response time to other bidders' actions makes eBay a second-price sealed bid auction (Bajari and Hortascu 2003). In such auctions, Vickrey (1961) proved that truth-telling is a dominant strategy. Thus, the eBay auction in this stage resembles a second-price sealed

bid auction, under which bidder i will have the incentive to bid $b_i(v_i) = v_i$. Thus consumer surplus cs^{ipv} accrued to the winning bidder of an auction under the independent private value setting is:

$$cs^{ipv} = b(v_{(1)}) - p = b_{max} - p \qquad (2)$$

where b_{max} is the highest winning bid. Note that b_{max} is not directly observable on eBay, but it is available to us from our Cniper.com.

I. B.II Case – Pure Common Value Setting

While it is commonly agreed that most auctions on eBay have elements of both private and common value components, Bajari and Hortascu (2003) point out that current analytical work is yet to determine equilibrium bidding strategies under this complex informational setting. Thus, like them, we consider the case of the pure common values setting as the alternative to the private value setting. In the context of equation (1) above, this is obtained by setting $\alpha = \beta$. The equilibrium bidding strategy for second price sealed bid auctions under the pure common value setting have been derived by Paul R. Milgrom and Robert J. Weber (1982). The primary consideration here for bidders is to avoid the winner's curse by shading their bids in an increasing fashion with the number of competing bidders. The pure common value or the "mineral rights" setting implies $v_i=v$, the ex post common value, and that v is same across all the bidders, but the signals x_i are interdependent. Only after conditioning on v (i.e., $x_i|V = v$) do they become independent. To estimate surplus we need to estimate v, since we already know the price p. Observe that under the common value setting

$$V = \alpha\left(x_i + \sum_{j \neq i} x_j\right). \qquad (3)$$

In a second price sealed bid auction, a bidder with signal x_i will be willing to pay anything up to her expected value conditional on her winning the object but being tied with just one other bidder with the same signal. Thus, the bid function for bidder i is to bid

$$b_i(v_i) = E(V \mid x_i) = \alpha x_i + \alpha \sum_{j \neq i} E(x_j \mid x_i) \tag{4}$$

Following Klemperer (1999), we assume that signals are uniformly distributed $X_i \sim U(0, 2V)$, and therefore the conditional distribution of a signal, given that it is below x_i (i.e., $x_j \mid x_j < x_i$) is $U(0, x_i)$. This implies $E(x_j \mid x_j < x_i) = x_i / 2$ which leads to the following bid function:

$$b_i(v_i) = E(V \mid x_i) = \alpha x_i + \alpha x_i + \alpha(n-2)\frac{x_i}{2} = \alpha\left(\frac{n+2}{2}\right)x_i \tag{5}$$

Choosing $\alpha = 1/n$, where n is the number of bidders, reduces this to the "average model" formulation as in Goeree & Offerman (2003). For $\alpha = 1/n$, we have the bid function

$$b_i(v_i) = E(V \mid x_i) = \frac{n+2}{2n} \cdot x_i \tag{6}$$

Under the assumption $X_i \sim U(0, 2V)$ each signal is unbiased for estimating V. However, the sufficient statistic for estimating V is the highest value $X_{(n)}$, and the maximum likelihood estimator for V is a function only of $X_{(n)}$. We therefore need only recover the signal of the winner, which we have through our Cniper.com dataset. This is also very useful from a practical point of view because of the presence of spurious bids at the start of the auction and the fact that the auction turns into a sealed-bid second-price auction only during the last moments of the auction (Bajari and Hortacsu 2003). Using only the highest bid keeps us away from such problems.

In order to recover $x_{(n)}$ we use the inverse of the bid function from equation (6):

$$x_i = b_i(v_i)\frac{2n}{n+2} \tag{7}$$

Plugging in the highest bid b_{max} in place of $b_i(v_i)$ gives us an estimate for $x_{(n)}$. In the case of $n = 2$ bidders, according to this formulation (as well as the classical formulation) each bidder assumes that the other bidder has an equal signal ($x_1 = x_2 = x$), and therefore $v = x$. In such a case each bidder should bid his/her signal, which results in no winner's curse. In our surplus computations for $n = 2$ we therefore use the difference between the highest bid and the price, similar to the private value setting.

9

Finally, the maximum likelihood estimator for V based on $X_i \sim U(0, 2V)$ is given by $x_{(n)}/2$. Notice though that this estimator is biased:

$$E(X_{(n)}/2) = \frac{n}{n+1} \cdot V \qquad (8)$$

A bias-corrected estimator for the common value (using (7)-(8)) is:

$$\hat{V} = \frac{n+1}{n} \cdot \frac{X_{(n)}}{2} = \frac{n+1}{n+2} \cdot b_{max} \qquad (9)$$

Thus, consumer surplus cs^{cv} in common value settings can be estimated by

$$cs^{cv} = \hat{V} - p \qquad (10)$$

Note that (10) allows for negative values of surplus, which in the context of common values would signal the occurrence of the winner's curse.

I. C Generalizability

Because our dataset arises from a sniping website, it raises questions about the generalizability of our results to the overall population of eBay auctions, where sniping may or may not occur. We address this empirically by testing whether a randomly drawn validation sample of 1000 eBay auctions has similar distributions of key auction parameters as do our field experiment data. We find that in all auction parameters, including item price, bidder and seller reputations, item categories, number of bidders and opening bid, there is no significant difference between our validation and field data[7]. This, coupled with the academic and practitioner support of the notion that sniping is widespread, leads us to believe that our findings are generalizable across eBay auctions.

I. D Summary of Findings

A simple analysis of the data shows that surplus is extremely right-skewed which makes the sample mean an unreliable measure of central tendency. Looking for a more robust and conservative measure, we find the sample median of surplus, i.e. the physical center of the surplus distribution, equal to

---

[7] Appendix A provides the comparative box-plots, QQ-plots, and additional charts indicating no significant difference in any of the variables.

$3.61. Our results indicate that consumer surplus in online auctions can be well-approximated by a 3-parameter Weibull distribution. We use the Weibull distribution to evaluate the sampling error of median surplus and show that a 95% confidence interval for the median is [$3.28, $3.73]. We estimate total surplus, based on an estimated 417.5 million[8] auctions which resulted in a sale on eBay in 2003, and find $1.5 billion accrued consumer surplus for the year 2003. On a percentage basis, consumers are capturing at least 18.3%[9] of the total surplus generated in eBay auctions. This high and growing consumer surplus level is one reason why online auctions are an attractive retail channel for consumers.

We find that US currency auctions carry higher surpluses relative to Euro and GPB auctions, by a factor of approximately 15%. Surplus levels in Euro and GBP auctions are similar to each other. There appear to be three main groups of surplus categories. The highest surplus is accrued to the group of eBay categories that are antique or collectible in nature. This is followed by a moderate surplus group of items comprised of computers, electronics and books, among others. The lowest surplus is in the group of household items such as toys, health and beauty items and games. We find that sellers with higher feedback ratings, a proxy for experience and trust, tend to yield lower bidder surplus, and that experienced bidders tend to realize higher surplus. We find that the main effects of price (proxying for stake), opening bid, and number of bidders have a significant influence on surplus, but so do the interactions of stake with opening bid and stake with number of bidders. These main effects must therefore be interpreted cautiously. Interestingly, we find that surplus is positively associated with stake in auctions that see many competing bidders, but this relationship is moderated by the opening bid. We find that surplus is generally increasing in auction duration and decreasing in sniping time, but only for

---

[8] eBay's 10k statement available at http://sec.gov/Archives/edgar/data/1065088/000089161804000676/0000891618-04-000676-index.htm has a section on operational parameters. They report that a total of 971 million items were listed in 2003. However, a significant percentage of auctions don't get any bids. Our conversation with a senior eBay executive revealed that eBay treats the overall success rate as confidential information. Thus, to estimate our multiplier we relied on aggregate level secondary eBay data from Hammertap.com . Based on a sample of 14,000 auctions from Hammertap we conservatively find an overall success rate of about 43 percent (lower 95% CI). Thus, we use 417.5 million as our multiplier. Further research is needed on analyzing the level and determinants of auction success rates on eBay.

[9] Based on median surplus and price levels.

"mainstream auctions" with five to seven day duration and sniping time equal to eight [10]or nine seconds.

We elaborate on these findings and on how we reach them in the remainder of the paper.

---

[10] The default on cniper.com

I.E Limitations

There are some limitations to our study that we would like to point out. Our model does not consider the presence of simultaneous auctions taking place on eBay. As pointed out by Zeithammer (2005) and Snir (2005), if bidders were to factor the simultaneous nature of auctions into their equilibrium bidding strategies, then bid shading would occur, which would deem our current surplus estimates as conservative. In addition, our wide-ranging data from a variety of eBay categories does not allow for easy access to outside "book" values, or consideration of what is going on in non-eBay markets. If indeed bidders use outside price comparison sites to cap their willingness to pay on eBay, then once again our consumer surplus measures are conservative. We also do not consider the time costs associated with using and submitting a bid on eBay, which are likely to be incurred by both winning and losing bidders. Lastly, because of the sniping nature of the data we do not consider auctions that close earlier due the exercise of the buy-it-now option.

The next section of this paper describes the working of the bidding agent and the data sample. Section III characterizes the distribution of consumer surplus, provides estimates of levels of consumer surplus in eBay auctions and explores the determinants of consumer surplus. Section IV concludes by pointing out limitations of the study and directions for future research.

II. DATA

II.A Description of the Bidding Agent

The prevalence of sniping on eBay has lead to several independent third party sniping agents that help bidders place last second bids on eBay. The interested reader is referred to Bapna (2003) for a review of sniping agents and their technical details. This study utilizes data from one such agent, Cniper.com, deployed by us. Cniper's logo "Snipe bids in your sleep, for free" is all explaining. Cniper is deployed just two hops away from eBay's server, making bid submissions lightning fast[11]. While most competing

---

[11] Roth and Ockenfels (2002) emphasize the probability of bids not getting through in the last seconds.

eBay sniping agents are fee based[12], Cniper has always been a free service and has a loyal and steadily growing user base of 2,035 bidders. It relies solely on word of mouth for advertising. In the period beginning July 23, 2003[13] and ending June 24, 2004, Cniper placed 69,571 bids on eBay on behalf of its users. Cniper is developed using PHP[14] and MySQL[15] and is deployed on an Apache webserver sitting on a Unix box. Thus, Cniper leverages the latest advances in open source software technology to keep its costs low[16]. This helps us provide it as a free service, and provides no incentive for any bid shading to account for bidding agent commissions. The lack of commissions also attracts entry for the tool, which in turn provides us with continuously richer observations of real economic agents acting in real markets. We believe that our approach, a first in the research community, will serve as a model for researchers doing Internet based field experiments and so-called "action research".

While the full technical details of Cniper.com's working are beyond the scope of this paper, a brief overview of its usage is necessary to motivate its usefulness in measuring consumer surplus. Agents such as Cniper, allow bidders to reveal a) their willingness-to-pay for a specific item being auctioned on eBay, and b) the number of seconds before the close of the auction they want their bid to be placed on eBay. For illustrative purposes, we continue with our earlier example of the Nokia 6610 GSM phone auction which was sniped and won by eBay user 'kanchenjunga' using Cniper and whose bid history is displayed in Figure 1. We show the process of the bidder sniping and the actual winning bid placed in Figure 2. Figure 2 reveals that bidder 'kanchenjunga' fired $180 for the item and won the auction by outsniping bidder 'ray7748' by 3 seconds. Recall, from Figure 1, that the winning bid or price is $170. Thus, it is evident that bidder 'kanchengjuna' derived a surplus of $10 from this auction.

---

[12] BidSlammer.com, AuctionSniper.com etc.
[13] This was when the site was significantly redesigned.
[14] A fast growing server-side scripting tool.
[15] The standard open source relational database.
[16] Cniper has zero licensing fees costs. Its only costs are those of hosting and the first author's time. The latter is bursty and can be significant at times when eBay changes its bid acceptance technology and Cniper has to respond by reprogramming its bid submission protocol.
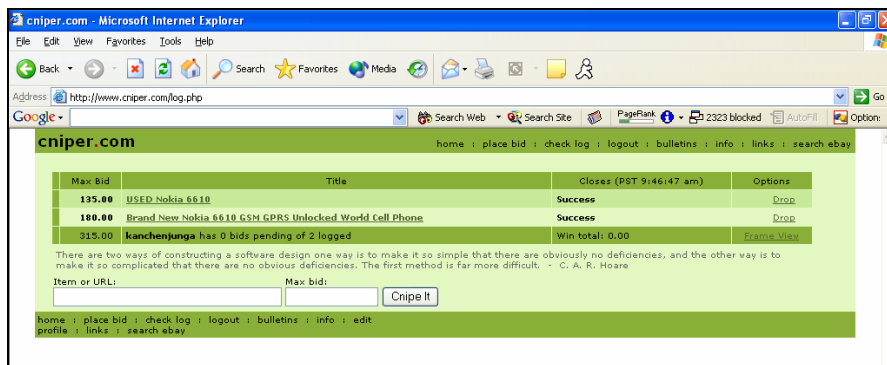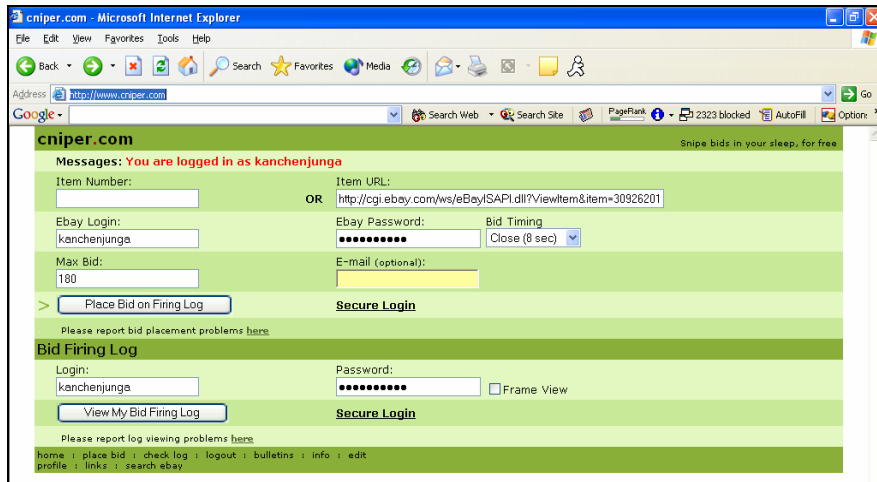
Figure 2 Top panel: Bidder *kanchenjunga* requests Cniper to bid $180 for eBay Item 092620119 eight seconds before the auction closes. Bottom panel: Row two of *kanchenjunga's* Cniper log shows the $180 bid as fired and successful

II. B. Summary Statistics

The data used in our analysis consist of 4514 eBay auctions that took place between January 9, 2004 and April 21, 2004[17]. In all these auctions the winner was a Cniper.com user. The auctions in our data were carried out in one of three major currencies: US Dollar (USD), Great Britain Pound (GBP), and the Euro. The items auctioned were in a wide variety of categories, spanning most[18] of eBay's 30 high

---

[17] This corresponds to little more than a three month period, the duration for which eBay posts bid histories of completed auctions.

[18] Only the two categories "Travel" and "Tickets" were not populated in our dataset.

level categories[19]. To maintain a minimal cardinality level in each category, we grouped the items, using eBay's categories, into 18 major categories. An additional 19th category was created for items in which the category description was missing[20].

To the best of our knowledge, currency and category have not featured in the extant analysis of eBay data. In addition, we recorded from eBay the following information on each auction: Opening and closing prices in their original currency and their USD equivalent[21], whether hidden reserve was used, the starting and ending time[22] and date, the number of bids placed in the auction, the number of unique bidders participating in the auction, and seller and winner rating. From Cniper.com we obtained the number of seconds before the auction close that the winning bid was placed, and the winning bid itself. Table 1 describes summary statistics for each of the variables in our study.

	Mean	StDev	Min	Max	Median
Winner Rating (W_RATING)	229.40	390.29	-3.00	11350.00	106.00
Snipe Time in Seconds (SNIPE_TIME)	8.69	0.90	1.00	13.00	9.00
Number of Bids	6.72	5.46	2.00	50.00	5.00
Number of Bidders (NUM_BIDDERS)	4.26	2.55	2.00	29.00	3.00
Seller Rating (S_RATING)	2588.15	9048.57	-1.00	170889.00	339.50
Auction Duration (NUM_DAYS)	6.66	2.10	1.00	10.00	7.00
Opening Bid (OPENING_BID)	19.29	101.68	0.01	3570.50	3.00
Price (PRICE)	60.78	229.39	0.08	7600.00	14.82
Surplus (PV)	15.59	56.12	0.00	1117.48	3.61
Surplus (CV)	5.15	47.43	-767.11	1019.02	0.18

	No	Yes
US Currency? (US_DOLLARS)	1782	2732
Hidden Reserve price?	4334	180

Table 1 – Summary Statistics of All Variables

---

[19] See http://pages.ebay.com/categorychanges/ for a list of high-level eBay categories. Notice that it includes the category "Everything Else" that contains auctions that do not fit any other classification. Also notice that the category "Automotive" is not contained in this list.

[20] While auctions with missing category descriptions could have been assigned to the "Everything Else" category, we decided to keep these auctions separate in order to maintain objectivity.

[21] eBay provides approximate conversions on the web page.

[22] Since eBay auctions last either 1,3,5,7 or 10 days, the starting and ending times are always equal.

II.C. Test for Private vs. Common Value

We are challenged by the diversity of our eBay auction data that ranges from pure common value, to part-private/part-common, to pure private value items. It is well known that private and common value settings lead to very different bidding patterns and correspondingly different auction outcomes [Milgrom and Weber (1982)]. Bajari and Hortascu (2003, page 337) comment on the fact that while the "part-private/part-common" value setting is perhaps the most realistic, the "analysis and structural estimation of such a model … is not tractable using the currently available methods." Thus, they use a winner's curse test to determine if their data (of coins) leans more towards a private or common value. Upon regressing the normalized bids with the number of bidders, while controlling for other auction factors, they find a significant negative relationship between bid levels and number of competing bidders. This indicates that, in their coin data, as competition increases, bidders shade their bids in increasing magnitudes to avoid the potential of incurring a higher winner's curse. They take this as "suggestive, but not conclusive," evidence of the presence of a common value setting, and for reasons outlined above, pursue a pure common value analysis.

We apply a similar test to our rather large and diverse dataset, and find no evidence for a negative correlation between the normalized bids[23] and number of bidders. In fact, the coefficient estimated from the model is +0.04 and highly significant (P-Value < 10-16). This relationship is also confirmed by scatterplots and other supportive analyses. There are other behavioral reasons as well that seem to suggest the private values setting. For instance, on Cniper.com bidders do have an option of revising their bids up until the last minute of the auction. They might be inclined to do this if they were getting influenced by other bidder's bids, as would be expected in a common value setting. Yet, our data suggest that the overwhelming majority of bids (95%) do not get revised.

Thus, following the approach of Bajari and Hortascu (2003), for the remainder of this paper we take this as suggestive, but not conclusive, evidence of a tilt towards a private value setting in our data. We recognize that while not perfect, this is a reasonable starting point for investigating consumer surplus

---

[23] Following Song (2004) bids were normalized using the third highest bid as proxy for the item's outside value.

levels and the important bidder, seller, and market characteristics that influence consumer surplus levels.

Clearly, more research is needed in understanding the weights of α and β at an auction level which would

indicate the degree of private/common value component.

We next estimate total surplus, analyze the distribution of surplus levels, and explore its

determinants.

III. ANALYSIS OF CONSUMER SURPLUS

III.A Estimating Total Surplus

We begin by estimating total surplus using our sample. Following the standard approach and for

sake of completeness, we begin by using the overall mean surplus estimate. Subsequently, we drill down

into category level means, finally gravitating towards a more robust median based analysis. A standard,

albeit crude, approach would be to use the overall mean private-value surplus multiplied by the estimated

total number of 417.5 million eBay auctions in 2003 (see footnote 8). This gives an estimate of $6.5

billion in total consumer surplus in eBay in 2003. However, since mean surplus varies somewhat across

categories, we can obtain a slightly narrower confidence interval by using category-weighted totals.

Table 2 reports mean surplus estimates based on the private value (PV) computation of Section I. The

weights are the estimated number of auctions in each category (estimated from the percentages in our

sample). This yields the 95% mean-based confidence interval [$5.832 billion, $7.188 billion].

Category	Proportion of Auctions in Data	Estimated Total Number of Auctions *	Mean Surplus (PV)
Antique/Art/Crafts	2.35%	9,803,943	$ 35.31
Automotive	5.58%	23,307,488	$ 19.72
Books	5.85%	24,417,368	$ 4.22
Business/Industrial	2.19%	9,156,513	$ 15.37
Clothing/Accessories	3.61%	15,075,875	$ 6.70
Collectibles	11.17%	46,614,976	$ 24.47
Computer/Networking	5.47%	22,845,038	$ 16.62
Coins/Stamps	3.48%	14,520,935	$ 8.36
Everything Else	2.92%	12,208,684	$ 17.85
Consumer Electronics	4.54%	18,960,456	$ 22.52
Health/Beauty	1.84%	7,676,673	$ 3.98
Home/Garden	6.89%	28,764,400	$ 8.13
Jewelry	4.34%	18,128,046	$ 25.72
Music/Movie/Video Games	10.97%	45,782,565	$ 4.56
Missing	10.66%	44,487,705	$ 12.71
Pottery/Glass	1.04%	4,347,031	$ 18.72
Photography/Camera	3.35%	13,965,995	$ 23.77
Sporting Goods	3.74%	15,630,815	$ 15.56
Toys/Hobbies	10.01%	41,805,494	$ 21.83
Total	**100.00%**	**41,750,000**	**$ 15.59**

* Based on a grand total of 417.5 million auctions

Table 2 - **Breakdown of data by eBay Categories**

An even more precise estimate of the total surplus can be computed by breaking down the data by price, because mean surplus levels vary across different price ranges. This is evident in Table 3. Using a weighted-total calculation as earlier, this time weighted by price-range, we obtain a further precise mean-based 95% confidence interval as [$5.890 billion, $7.131 billion].

Price Range (endpoint)	% Auctions in Data	Estimated Total Number of Auctions	Mean Surplus (PV)	Median Relative Surplus (Surplus/Price)
$1.00	0.33%	1,387,350	$3.67	3.06
$1.65	1.00%	4,162,051	$2.83	1.26
$2.72	5.05%	21,087,727	$2.54	0.56
$4.48	9.95%	41,528,024	$3.12	0.43
$7.39	11.92%	49,759,637	$4.24	0.31
$12.18	15.18%	63,355,671	$5.42	0.25
$20.09	15.31%	63,910,611	$7.06	0.24
$33.12	12.89%	53,829,198	$11.14	0.18
$54.60	9.08%	37,920,913	$20.20	0.19
$90.02	6.27%	26,174,679	$23.70	0.14
$148.41	5.03%	20,995,237	$30.01	0.12
$244.69	3.59%	14,983,385	$51.82	0.10
$403.43	1.99%	8,324,103	$62.51	0.06
$665.14	0.95%	3,977,071	$101.67	0.08
$1,096.63	0.82%	3,422,131	$124.86	0.08
$8,103.08	0.04%	184,980	$84.53	0.09
Total	**100%**	**417,500,000**	**$15.59**	**0.22**

Table 3: Mean surplus and median relative surplus for items in different price ranges

Notice that although we use surplus means for computing total surplus, means are not reliable measures of an auction's "typical" surplus value, because surplus is extremely right-skewed and thus the mean does not describe the physical center of the surplus distribution well. Next, we derive a more robust and also more conservative approach.

III.B Surplus Distribution

Figure 3 shows a bar chart of surplus of the USD data; the other two currencies are similar. Two interesting observations are worth considering. Firstly, we see that the values $0, $0.50, $1.00, $1.50 and so on are over-prevalent, that is, more prevalent than values in between. Only a very small proportion of surpluses assume values between, say, $0 and $0.50, or between $1 and $1.5. For instance, while 75 of the surplus values equal exactly $1, only 10 equal $0.99, only 6 equal $0.98 and only 4 equal $0.97. On the other hand, only 32 values are $1.01. This is surprising, since we would expect surplus to be distributed uniformly in such small intervals. Thus, our first observation of the distribution of consumer surplus is that the data are apparently *semi*-continuous. By semi-continuous we mean that while surplus can theoretically assume any value in a given interval, some underlying (and unobserved) data-generating

mechanism introduces discretization, causing the data to be concentrated on certain values. We explore the possible causes of this later in this section.
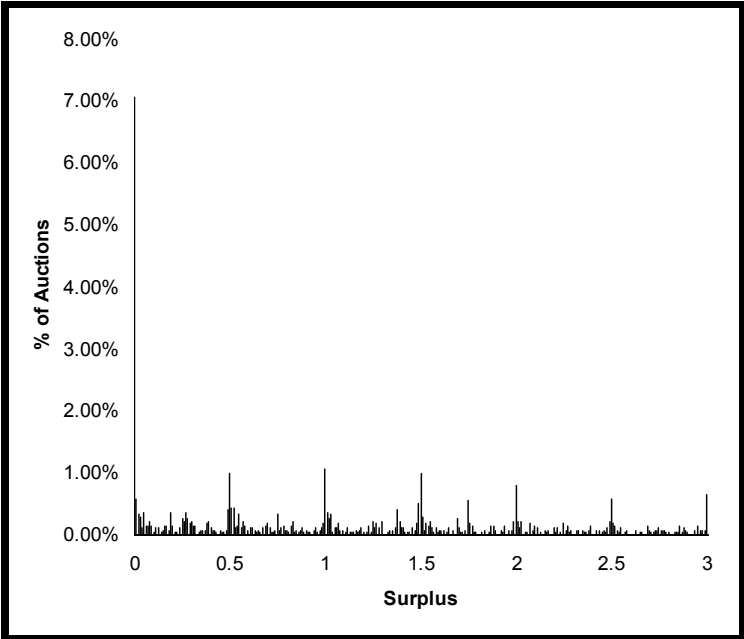

Figure 3 Bar chart for Surplus in the Range $0 to $3

In order to transform surplus from auctions in various currencies into a single scale, we converted all currencies into USD using the conversion rate listed on the auction's eBay page. Figure 4 displays a histogram of log(USD surplus+1). The shift of 1 allows us to apply the log transform to the zero surplus data. The data are clearly bi-modal, with a large "lump" of zero values (0= log(1)).
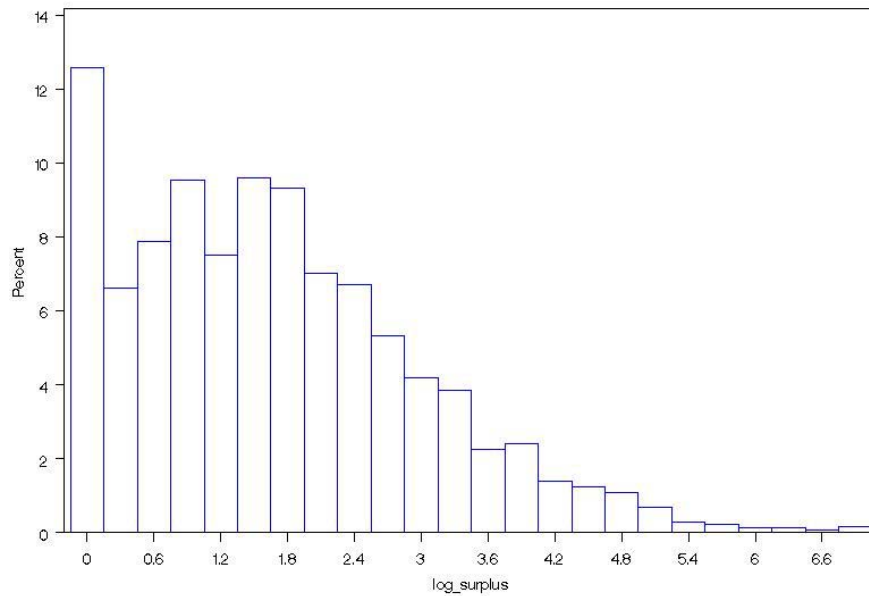
Figure 4– **Surplus Distribution is Bi-Modal**

Our second observation is that 368 (out of 4514) auctions ended at a closing price exactly equal to the revealed willingness to pay indicated by the winner to Cniper. These can be categorized as highly competitive zero surplus auctions where the two highest valuations are identical. Given that our research objective is to quantify and understand the determinant of consumer surplus, the above "zero-inflated" data present a challenge, in that we cannot directly model surplus in a regression model without violating the model assumptions.

III.B.I Finding a Suitable Transformation

Although there are multiple values with high frequencies, zero-inflation is the most prominent. In order to model such data we can take two approaches: We can transform the data to achieve a more regular distribution, or develop a model that can account for such a data structure. An example would be a mixture model that results from combining a process that generates only special values such as zeros with another process that generates non-negative or positive surpluses. The first approach, of transforming the data, is more popular because it is simpler to use and understand. It requires fewer assumptions and fewer parameters, and enables to make use of a large variety of established statistical methods. The only

challenge is to find a good transformation. It turns out that ordinary transformations do not yield satisfactory results. For instance, the Box-Cox transformation indicates that the log transformation brings the data closest to normality. However this does not alleviate the heavy zero-inflation and the resulting data are clearly not normal. Furthermore, we would like the transformation to deal with the other (non-zero) high-frequency values as well. Our solution is to use a novel transformation which moves the data into a slightly coarser scale. In particular, we round surplus to the next integer value. The reasoning behind this transformation is data driven: Recall the initial observation about the semi-continuous nature of the data, with large counts of surpluses at values of 0, 0.01, 0.5, 1, etc. We believe that the reason is most likely the discrete bid increments that eBay imposes[24]. Indeed, according to eBay's rules, when the difference between the two highest bids is less than an increment, the winner pays the price that s/he bid, resulting in a zero surplus. If this is contributing to the zero inflation, then we would expect most of the zeros to be coming from auctions with high closing prices, where the bid increment is large and therefore the chance of the highest two bids being closer than an increment is higher. To check this we compared the distribution of closing prices for zero-surplus auctions vs. non-zero-surplus auctions. We find that while the price distribution is very similar in both groups (the median is nearly identical), there do seem to be more high-price auctions that closed with zero-surplus than non-zero surplus. This indeed could explain the inflation in zeros. However, there are also very large masses at several other values (such as $0.50, $1.00), indicating that there is another mechanism contributing to the "special value" phenomenon. It could also be the case that most bidders, unsure whether their valuation of an item equals, say, $9.95 or rather $10.00, enter integer values as their bids, and this further contributes to semi-continuous nature of the data. Reasoning that surpluses within a $1 unit range carry the same (or at least similar) information, we transform the original surplus data by applying the ceiling function (which gives the next highest integer). We then take log(integer surplus + 1) to accommodate for the zero surplus data. Figure 5 shows the histogram of the transformed integer surplus values. The new values follow a unimodal, right-skewed distribution.

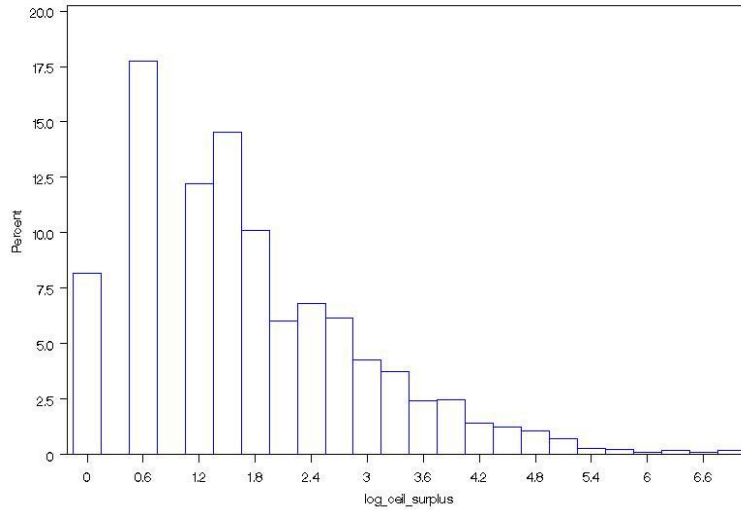[24] See Bapna et al (2003b) for the impact of discrete bid increments on online auctions

Figure 5– (log) Ceiling transformed Surplus is right-skewed

Further investigation of the probabilistic structure of the data reveals that a 3-parameter Weibull distribution approximates the transformed data fairly well. This can be seen in Figure 6 which displays a Weibull probability plot of the data (left), and in comparison a lognormal probability plot of the same data (right).
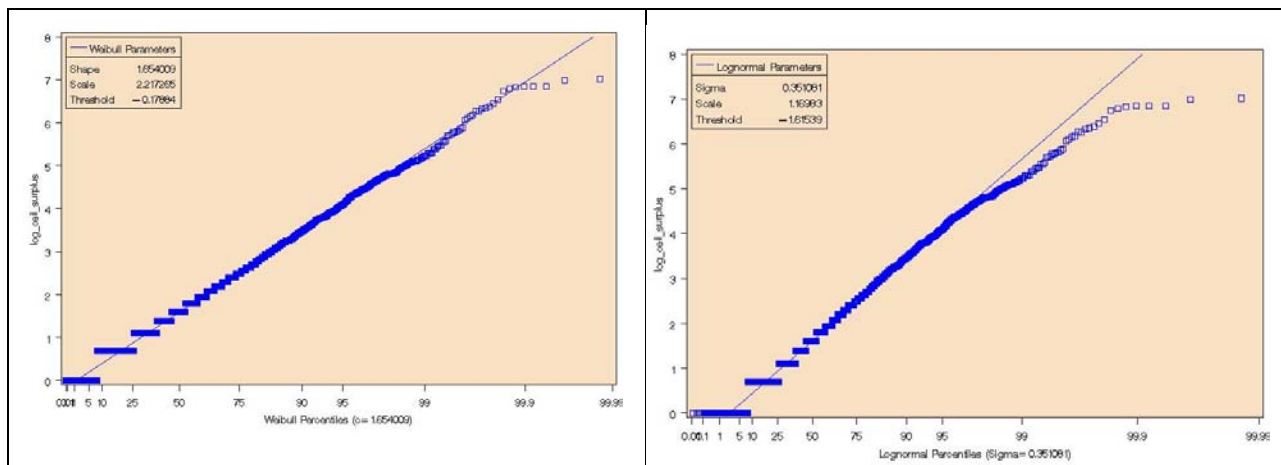

Figure 6 – Three parameter Weibull distribution fits better than a Lognormal distribution (right)

This approximation is of special interest, since the Weibull distribution often describes the time to an occurrence of the "weakest link" in competing failure processes; for example, the time to death of the first component in a system with multiple components and competing failures. In our case we have a set

24

of bids in an auction. Among the differences between each of the bids and the winning price, the surplus is the only non-positive value, and thus the minimum of these differences. In this sense the winning/highest bid is the "weakest link" because it has the smallest distance from price compared to all other bids that were placed in the auction. Taking logs maintains the order, and therefore the result holds. It is also interesting to note that the relative surplus (the rounded surplus divided by price) also follows a Weibull distribution.

The estimated parameters of the fitted Weibull distribution are shape = 1.65, scale = 2.22, and location/threshold = -0.18. Since the surplus distribution is very skewed, the median is a better measure than the mean for the center of the distribution. Although we can compute the median surplus of $3.61 directly from the original values (converted to USD), we also prefer having a measure of sampling error in order to quantify the uncertainty in surplus generated by a "typical" ebay auction. We use the asymptotic normality [Herbert Aron David and Haikady Navada Nagaraja, (2003), page 241] of the sample median and the Weibull distribution to obtain the formula for the median's standard deviation. After transforming the data back into its original units we obtain a 95% confidence interval of [$3.28, $3.73][25]. We also obtain a non-parametric confidence interval by taking bootstrap samples from the raw (i.e. un-transformed) surplus data. A 95% bootstrap confidence intervals (based on 10000 replications) yields [$3.50, $3.88]. The similarity of these intervals indicates that our data-transformation maintains the main features of the data very well.

The median is also a useful statistic for quantifying the lower bound on percent surplus: Following Giray, Hasker and Sickles (2005) we use a ratio based on the medians of price and of surplus to derive the estimate of the lower bound of % surplus. We estimate this by

$$\underset{i}{Median}\left(\frac{surplus_i}{price_i + surplus_i} \right) = 18.31\% \ .$$

---

[25] The median is asymptotically normal, centered around the population median and with variance $[\ 4\, n\, f^2(Med)\ ]^{-1}$, where $f(Med)$ is the Weibull density at the median [David and Nagaraja (2003), page 241]. We estimate this quantity using the estimated Weibull parameters. This yields a standard deviation of 0.023 for the median of log(integer-surplus). A 95% CI for the population median of log(integer-surplus) is [1.56, 1.65]. Taking an exponent and subtracting 0.5 (the average rounding) yields the 95% CI for median surplus of [$3.28, $3.73].

III.C. Determinants of Consumer Surplus

Prior empirical research dealing with consumer surplus in online auctions has made relative comparisons of groups of bidders in discriminatory multi-unit Yankee auctions [Bapna et al 2003a, 2003b, Bapna et al. (2004)]. Their main finding is that *evaluators*, bidders who make singleton early high bids, leave significantly more money on the table in comparison to other groups of bidders. To the best of our knowledge, there has been no other study that has looked at determinants of absolute consumer surplus levels.

In this paper, we work with the ceiling transformed surplus data and identify candidate explanatory variables that could influence the level of bidder surplus. It should be noted that much of the prior research has focused on understanding the determinants of auction price. Bidder surplus, of course, is intrinsically associated with price, albeit negatively. We believe that the following five categories of independent variables potentially influence the rents that bidders can extract.

1) *Seller's mechanism design choices***:** Sellers, who strive to maximize their revenues, can be expected to strategize on eBay by choosing the appropriate combination of opening bid level, auction duration and the usage of a hidden reserve price. Opening bid can be interpreted as an open reserve price and prior research has contrasted the comparative effectiveness of open versus hidden reserve prices on sellers' expected revenue. Myerson (1981) formulates the optimal (seller revenue maximizing) auction design problem as being equivalent to deriving the optimal open reserve price. Rama Katkar and Lucking-Reiley (2000) in a field experiment selling Pokemon cards, find that hidden (secret) reserve prices make sellers worse off, by reducing the probability of the auction resulting in a sale, deterring serious bidders from entering the auction, and lowering the expected transaction price of the auction. In contrast, Bajari and Hortaçsu (2003) based on an econometric estimation, suggest that optimally chosen hidden reserve prices can yield the seller one percent higher revenues. Thus, the evidence seems mixed with respect to how the seller's usage of hidden reserve prices impacts the auction price, and consequently the bidders' surplus.

It is well established that, on eBay, lowering opening bids attracts more bidders [Bajari and Hortaçsu (2003), Lucking-Reiley et al. (2000)]. In addition, prior research also suggests that when a seller chooses to have her auction last for a longer period of days, this significantly increases the average auction price [Lucking-Reiley, Bryan, Prasad, and Reeves (2000)]. We describe the number of competing bidders under market characteristics and examine its interaction with opening bid. Taken together, the prior research[26] allows us to hypothesize:

H1: Opening bid has a negative influence on bidder surplus

H2: Auction duration has a negative influence on bidder surplus

2) *Seller characteristics*: eBay's feedback reputation system has been widely studied and many studies indicate that sellers with higher reputations engender trust and extract premiums [Sulin Ba and Paul A. Pavlou (2002)]. It can also be argued that sellers with more experience[27], i.e. higher seller rating, draw more buyer participation, which can reduce the difference between the highest and the second highest bid. Thus, we hypothesize:

H3: Seller reputation has a negative influence on bidder surplus

3) *Product characteristics*: For homogenous products, by definition surplus increases with a decrease in price. However, in contrast to the above-mentioned empirical studies on eBay [Bajari and Hortaçsu (2003), Katkar and Lucking-Reiley (2000), Lucking-Reiley et al. (2000)] that controlled for product heterogeneity, our dataset is diverse, covering all but two of eBay's 30 major item categories, with prices ranging from 1 cent to $7600. This allows us to test the implications of stakes and product attributes on surplus levels in a far more generalizable setting. Smith and Walker (1993) have predicted that individuals' behavior will more closely match the predictions of rational behavior as the stakes of the

---

[26] Prior research typically considers only main effects and our hypotheses reflect this. As our detailed analysis will show though, some of the main effects are moderated by interaction terms.

[27] eBay's feedback rating, which indicate the difference between positive and negative ratings, have been viewed as indicators of experience as they are generally reflective of the number of transactions conducted by the seller. Paul Resnick and Richard Zeckhauser (2001) report that only 0.6% of feedback comments left on eBay by buyers about sellers was negative or neutral.

decision increase[28]. Marketing theories suggest that as stakes get higher, consumers get more involved in finding the best price for their product [Marcel Cohen (2000)]. Based on the mixed evidence regarding the influence of the item's stake and in light of our diverse dataset we do not specify a directional hypothesis. Instead, we hypothesize:

H4: The item's stake magnitude has a significant influence on bidder surplus

4) *Winner characteristics*: We collect information on the winning bidder's rating as a proxy for their experience on eBay. We expect more experienced bidders to have more strategic ability and possibly confidence in their valuations and bids. Hence we expect them to derive higher surplus. In addition, we also measure aggressiveness of the bidder by looking at how many seconds prior to the close of the auction they snipe. Given that there is a small likelihood of bids not getting through due to congestion on eBay, even when fired by agents such as Cniper[29], we associate risk seeking behavior with a greater desire to win and consequently higher bid levels and surplus. Thus, we hypothesize:

H5: Winner experience has a positive influence on bidder surplus

H6: Winner aggressiveness has a positive influence on bidder surplus

5) *Market characteristics*: We are fortunate to have significant data in three prominent currencies, namely USD, GBP and the EURO. This allows us to test, for the first time, whether bidders differ in their bidding characteristics across countries. While there has been prior research in looking at the efficiency of auction formats across different countries [Klemperer (2002)], this study represents a first in comparing bidder surplus levels across countries. Given that eBay was founded in the US and subsequently expanded to UK and Europe, we hypothesize that:

H7: The longer experience of the US market has a positive influence on bidder surplus

Lastly, as can be expected in a private value setting, we propose that the level of competition in an auction, reflected in the number of bidders, negatively influences the surplus accrued to the winning bidder. We also have data on the total number of bids in an auction. However, the total number of bids is

---

[28] In the context of auctions this maps to surplus maximizing behavior.
[29] Cniper offers a choice of as low as 2 seconds before closing, but calls it "insane."

highly correlated with the total number of bidders and thus does not add any additional explanatory power to our regression models. Thus, we hypothesize that:

H8: Increasing competition has a negative influence on bidder surplus

III.D Modeling Approach

Our goal is to find a model that explains the effect of the above-mentioned set of explanatory variables on consumer surplus. The main challenge with surplus data such as ours is the extremely large number of auctions resulting in "special values" and especially zero-surplus. That is, auctions which ended exactly at the value placed by the winner. Recall that in our dataset, 368 auctions (from a total of 4514) carried zero surplus. Furthermore, these auctions are diverse in their other characteristics (across categories, currencies, etc.). Such data cannot be modeled directly by ordinary regression without violating the model assumptions, and, as a consequence, yielding potentially misleading results. As mentioned before, two possible solutions are to transform the data and then use ordinary statistical models or to develop a specialized model that captures this behavior. Although "zero-inflated" models exist for discrete data (e.g., "zero-inflated Poisson", Lambert, 1982), we are not aware of such models for continuous data, especially when the inflation is not only for zeros but also for other values that cannot be easily predetermined. In order to fit such a model, we need to specify the special values and their probabilities, the mixing distribution, and the general surplus distribution. This results in many parameters to be estimated. In addition, the mechanism that generates the special values is not observable, and thus specifying "special values" based on the sample values alone can be misleading. If we make a simplifying assumption that only zero-surplus auctions should be separated from positive-surplus auctions, then we could use a two-stage model for price as a function of our hypothesized determinants: The first stage finds variables that are useful for classifying an auction as a zero- or positive-surplus auction. This can be done by using a classification method, such as logistic regression. The second stage fits an ordinary regression model to positive-surplus auctions, in order to find variables that determine surplus. The results of fitting such a model to our data are given in Appendix B. While this two-stage

approach is valid, one of its disadvantages is the loss of statistical power. Also, it is less elegant than a single integrated model, requires the estimation of many more parameters, and results in more complex, harder to interpret relationships. We therefore strive for an alternative approach that combines zero-surplus and positive-surplus (including the other high-frequency special values) under one roof.

As described in Section II, the rounding of surplus values and taking a log transform yields a distribution that is suitable for regression modeling. We fit a regression model to the transformed data based on our hypotheses above and using model selection procedures. The best model is given in Table 4[30]. Note that among the candidate predictors identified in Section III.B the usage of the secret reserve price, the number of bids, and the auction's duration did not add additional explanatory power to the model. The model's R^2 and Adjusted R^2 values are 0.3048 and 0.3016, respectively. In comparison, the best model *without interaction terms* had corresponding values of 0.2931 and 0.2902. The stake variable, Price, has the highest explanatory power, resulting in an R^2 value of 0.2468 for the simple, single-variable regression model. It is also interesting to note that when using only the "standard" eBay explanatory variables,[31] opening Bid, winner and seller rating, number of bidders, duration of the auction and price, the best model fit amounts to only 0.1265 and 0.1256 for R^2 and Adjusted R^2.

Notice that the model above uses log(surplus) as the response variable and log(price) as a predictor. This is mathematically equivalent to a model where the response is log(surplus/price), or log *relative surplus*, with the predictor log(price), where the coefficient for log(price) is equal to β_{price} -1 (i.e., we are subtracting log(price) on both sides of the equation). The coefficients for other predictors are the same as in the first model. The practical difference is in the interpretation that now looks at the *relative surplus* as a function of the item's price. We describe our results on the relationship between surplus and price in both terms below.

---

[30] The qualitative results for this model are very similar to the results obtained from a two-stage model, however it has more power.

[31] We say standard since many other empirical studies do not consider currency or ebay categories.

Variable		Estimate	Std. Error	Pr(>|t|)
(Intercept)		2.9826	0.5120	0.0000
CATEGORIES *				
	Antique/Art	0.4384	0.1022	0.0000
	Pottery/Glass	0.3130	0.1502	0.0372
	Collectibles	0.3759	0.0513	0.0000
	Everything Else	0.3480	0.0915	0.0001
	Toys/Hobbies	-0.1370	0.0539	0.0110
	Music/Movie/Games	-0.2049	0.0536	0.0001
	Jewelry	0.3581	0.0781	0.0000
	Automotive	0.1578	0.0691	0.0225
	Home/Garden	-0.2010	0.0623	0.0013
	Health/Beauty	-0.2756	0.1144	0.0161
US DOLLARS **		0.1479	0.0346	0.0000
NUM_DAYS		-0.1656	0.0692	0.0167
SNIPE_TIME		-0.2275	0.0572	0.0001
NUM_BIDDERS ***		-0.5580	0.0715	0.0000
PRICE ***		0.3042	0.0327	0.0000
S_RATING ***		-0.0288	0.0084	0.0006
W_RATING ***		0.0299	0.0115	0.0091
OPENING BID ***		-0.1406	0.0231	0.0000
OPENING BID x PRICE		0.0298	0.0051	0.0000
PRICE x NUM_BIDDERS		0.0999	0.0188	0.0000
NUM_DAYS x SNIPE_TIME		0.0224	0.0079	0.0047

* Base Category: Books, Business/Industry, Clothing/Accessories, Computer, Coins/Stamps, Electronics, Photography, Sporting Goods
** Base category: Euros and GBP
*** The variables surplus, price, opening bid, winner rating, seller rating and number of bidders were transformed to the log-scale

Table 4 - Parameter Estimates for regression on log(Surplus)

IV. ANALYSIS OF RESULTS

We organize this section by reporting first on the main effects and subsequently on those variables that have significant moderating interactions. We begin with the effect of the categorical variables of currency and product categories.

Currency: Consumer surplus is not the same across different currencies. We see that US currency auctions carry a significantly higher surplus relative to that of EURO and GPB auctions. In fact the coefficient equals .15 which implies that this difference is exp(.15) = 1.16. The surplus accrued to US

bidders is about 16% higher than in the European market! Whether this is due the longer experience that US consumers have had with electronic markets, due to cultural reasons, or to other unobservable factors remains an unanswered and interesting research question.

Product Categories: Three main groups of categories determine the surplus level:

- *Highest Surplus*: Antiques/Art, Automotive, Collectibles, Everything Else, Jewelry and Pottery/Glass have the highest surplus by an estimated factor of 1.3-1.5 relative to the reference group.

- *Moderate Surplus*: Computers, Electronics, Books, Clothing/Accessories, Coins/Stamps, Photography, Sporting Goods, Business & Industry and missing category descriptions. These categories comprise the reference group.

- *Lowest Surplus:* Toys/Hobbies (TH), Music/Movies/Games (MG), Home/Garden (HG), and Health/Beauty (HB) have the lowest surplus, by an estimated factor of 0.77-0.87 relative to the reference group.

The highest surplus group appears to be made up of items, such as antiques and jewelry, which require expertise in assessing valuations and could potentially also have a common-value component. Thus, there is a potential winner's curse effect revealed here. The moderate and low surplus groups are made up of items that have several alternative channels, such as eBay's half.com and Amazon.com, and are therefore very competitive.

Next, we report on the effect of seller and winner experience, as reflected in their total feedback ratings on eBay.

Winner rating: More experienced winners are associated with a higher surplus. This indicates that with increasing experience bidders act more strategically, identifying wining bid levels and maximizing their utility from participating in eBay.

Seller rating: In contrast, and as expected, a seller's experience has a negative relationship with surplus. It can be expected that more experienced sellers make better mechanisms design choices. In addition, they can be expected to be more sophisticated in describing the product, by choosing a better layout and pictures to attract more bidders. This enables them to extract rents away from bidders. It is also likely that

bidders simply trust more in sellers with a better reputation and this translates to more confidence in bidding. Overall, this effect is consistent with previous research [Ba and Pavlou (2002)] that indicates that higher seller reputation results in a higher price! Interestingly, the lack of significant interactions between the seller experience and the mechanisms design variables, suggests that only the reputation effect of the seller is at work. More work in needed in isolating these effects.

We now report the main effects that are moderated by interaction terms amongst them.

Price, Number of Bidders, and Opening Bid: The main effect of the stake variable, price, is positive, while the main effect of number of bidders and opening bid is negative. It is interesting to note that all of these three main effects are statistically significant, but so are the interactions of price with opening bid and price with number of bidders. These main effects must therefore be interpreted cautiously. The relationship between price and opening bid and its effect on surplus is visible in Figure 7.

The left panel is a scatterplot of surplus vs. price, with color denoting opening bid (continuous grayscale; gray = low, black = high). Overall, surplus is positively correlated with price, but the extent of this relationship depends on the opening bid amount: It is stronger in auctions with higher opening bids and weaker in auctions with low opening bids.
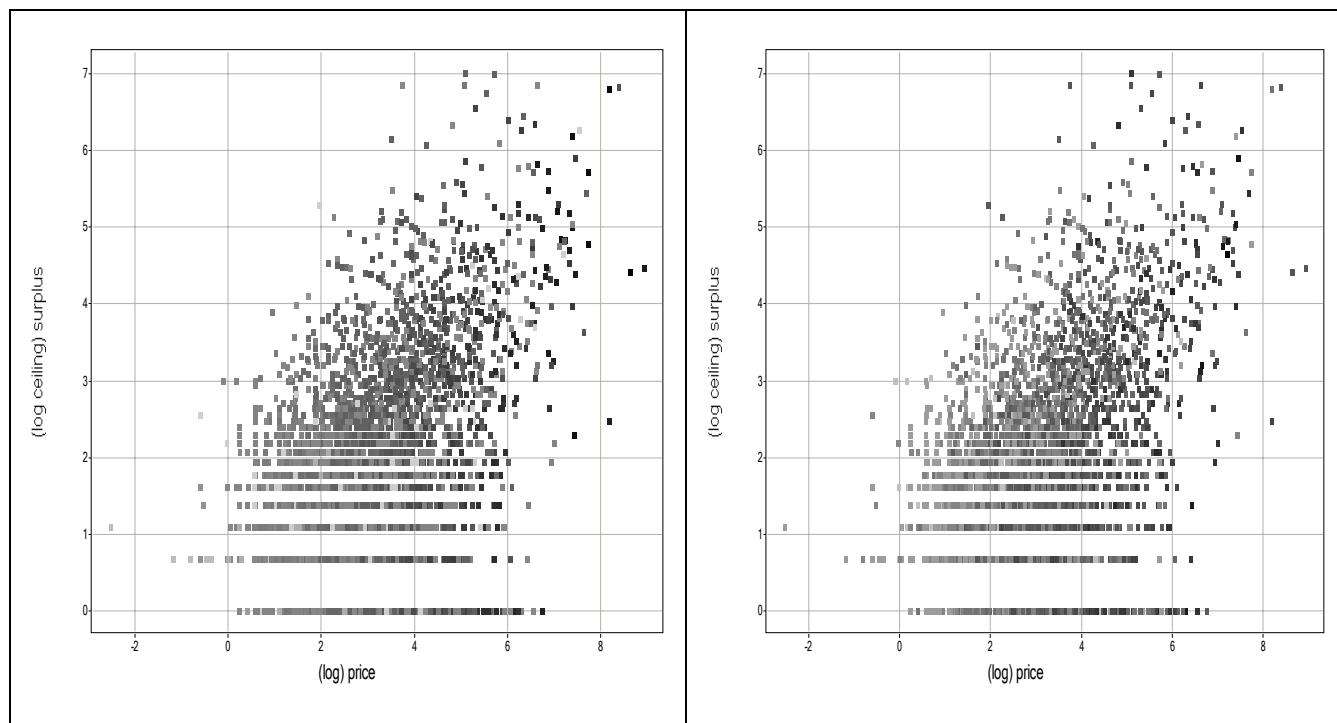
Figure 7– Surplus vs. Stake (all USD). Left panel: color = (log) Opening bid, High = black.

Right panel: color = (log) Number of bidders, Many=black.

The relationship between price and number of bidders and its effect on surplus can be seen in the right panel of Figure 7, which is a scatterplot of surplus vs. stake, with depth of grayscale denoting number of bidders (gray = few bidders, black = many bidders). The positive relationship between surplus and price appears to be stronger in auctions with many bidders, as can be seen by the slope of the cloud of dark gray to black auctions. In contrast, in auctions with few bidders, the relationship is weaker (gray color has lesser slope). The reason for this could be increased competition, which in turn causes the winning bidder to be less cautious, thereby generating higher surplus.

If we look at the relative surplus, log(surplus/price), as the response then the coefficient for price is negative (*0.30 -1 = -0.70*). Considering the number of bidders (between 2-29 in our data) and the range of opening bids ($0.01-$3570.50) we get a negative coefficient tying relative surplus with price. This means that higher value items accrue lower relative surpluses. Table 3 shows this effect by giving the median relative surplus for different price ranges. Notice that the interval lengths increase exponentially (reflecting the log scale).

To summarize, the regression model reported in Table 4 indicates a significant positive coefficient for the log-price (even when considering the interaction terms). Since the response is log-surplus it implies that a percentage increase in item price is associated with a percentage increase in surplus. The relationship of price with relative surplus, on the other hand, is negative. Interestingly eBay's gross merchandise sales grew by 160%, from $14.9 billion in 2002 to 23.8 billion in 2003. This suggests surplus should have gone up in the same period by an average of 49%, indicating growing benefits to consumers participating in online auctions. Our examination of the interaction effects reveals that the main effect of price has to be qualified by saying that overall surplus is positively associated with price, and this relation is stronger in auctions with higher opening bids. In addition, the rate of this surplus increase changes for auctions with different competition levels. Auctions with higher competition levels, in which there are larger number of bidders, see a faster surplus increase than auctions with a low level of competition.

The opening bid plays an interesting role in influencing surplus. Its main effect is negative, that is, higher opening bids are associated with lower surpluses. But also notice again the positive interaction effect with the stake variable, price: A higher opening bid results in a faster rate of surplus increase due to price. A closer inspection of the data reveals the explanation for this result. For low valuation items, setting the opening bid to something different than eBay's default (=$0.01) results in a negative correlation with surplus. On the other hand, high stake items have a positive correlation with the opening bid. This suggests perhaps that the signaling role of the opening bid varies with stake, a promising area of future research.

Duration and Sniping Time: Auction duration and sniping time (the number of seconds before the close of the auction that the bid fired), both have negative main effects. Overall, longer auctions and early firing times generate lower surplus. These results are as expected: Longer auctions result in higher prices and thus in a lower surplus. Very late firing times appear to be indicative of confident and perhaps risk-seeking bidding behavior.

However, the significant interaction between the two effects, which might be a result of the discrete nature of auction duration (=1,3,5,7,10 with high concentration at 7,10) and fire time (=2,4,8,9,10,..24 with high concentration at 8-9) tell us a more complicated story. The scatterplot in Figure 8 plots sniping time vs. duration with size representing surplus. It is apparent that surplus increases in auction duration for sniping time equal to eight and nine seconds but not for other values. It also appears that in five and seven day auctions surplus decreases for earlier firing times, but no such relation appears for other durations. In summary, surplus is generally positive in duration and negative in sniping time, but only for "mainstream auctions" with five/seven day duration and sniping time equal to eight [32]or nine seconds.
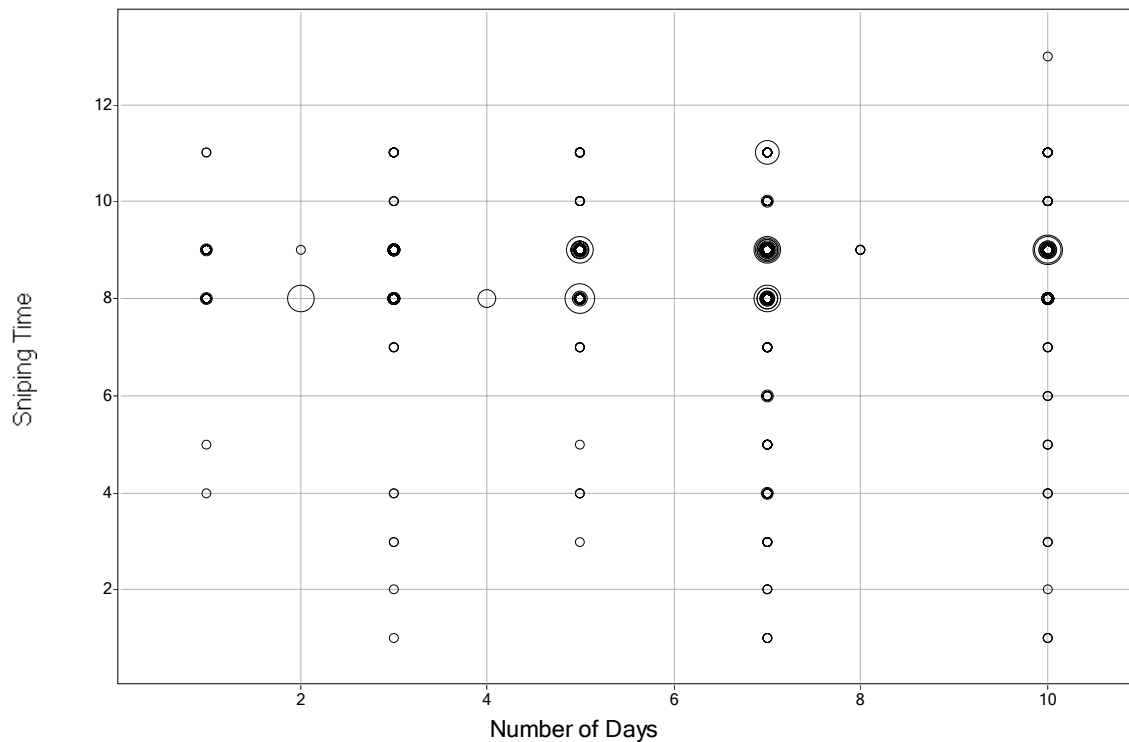
Figure 8– S**niping time v. Auction Duration (size proportional to Surplus)**

---

[32] The default on cniper.com

V. CONCLUDING REMARKS and FUTURE RESEARCH

eBay auctions are at the forefront of e-commerce, demonstrating how the Internet can remove spatial and temporal constraints to make economic exchange mechanisms, such as auctions, mainstream. In 2003, sellers through eBay sold $23.8 billion worth of merchandise. While gains from such trades to sellers and eBay, the market-maker, are obvious, little is known about consumer welfare levels accrued to buyers in electronic markets. We attempt to fill this gap by developing consumer surplus estimation procedures for common and private value settings on eBay. A key ingredient in both settings is the highest bid level that is unobservable on eBay, but available through an ongoing field experiment that allows eBay bidders to use a web based sniping agent called Cniper.com for free. In return, roughly 2000 bidders who regularly use Cniper.com to place last second bids on eBay, vote with their dollars and provide us with a wealth of information about actual willingness to pay for a wide variety of items sold on eBay. Keeping the service free attracts entry for the bidding agent, and provides no incentive for any bid shading to account for bidding agent commissions.

Our analysis of winning bids from 4514 auctions spanning all but two of eBay's categories reveals interesting insights into the distribution of consumer surplus. We find that eBay's discrete bid increments tend to make surplus data multi-modal and semi-continuous, rather than continuous. In addition, we observe zero-inflation (383 auctions have zero surplus), indicating a high degree of competitiveness on eBay. For an auction to have zero surplus, the two highest bids have to be exactly equal. These issues, coupled with data in three currencies and multiple categories catalyze us to innovatively transform our data using a log-ceiling transformation to make it suitable for modeling using standard statistical methods. We expect that our finding that a 3-parameter Weibull distribution best typifies consumer surplus to be useful to future researchers making distributional assumptions for analytical models examining the dynamics of online auctions.

We find that the stake variable, price, of the auction has the biggest positive influence on the bidder's surplus. This reinforces the belief that stakes matter, and as stakes get higher rationality (surplus maximizing behavior) becomes more prevalent. In addition we find that surplus levels are highest for

items such as collectibles and antiques, and US bidders accrue significantly higher surplus than their European counterparts. The 16% difference between the more mature U.S. and non-U.S. auctions raises interesting propositions regarding the long-term trends we can expect with consumer surplus. In particular, future research can be expected to shed more light on which parties gain as markets mature. How these longitudinal patterns will shape out and how they will impact buyer and seller entry promise to remain interesting questions for future research. It also points towards potential directions for future research that may investigate cultural differences in bidding behavior.

We find that sellers' mechanism design choices, namely the choice of opening bids, auction duration as well as the sellers experience all have a significant influence on surplus. So do bidder characteristics such as experience and aggressiveness, as indicated by how close to the auction's end they select to fire their bid.

While we attempt to check the generalizability of our results by examining potential distributional differences in any of the independent variables between our field data and a randomly selected test data of 1000 auctions, we realize that this test data set is only a small fraction of the auctions conducted by eBay on a given date. We are limited in this aspect by having only the public access to eBay's search engine, which is clearly not designed for such a purpose.

We believe that our findings that consumers are capturing at least 18.3% of the total surplus generated in eBay auctions is conservative. This belief stems from our conservative estimate of the number of transactions completed on eBay and also from the fact that we do not consider (or have estimates of) the auctioneer's and the seller's surplus in this calculation. Both these would have negative signs in the denominator of such a calculation.

As in traditional consumer theory, having benchmark consumer welfare levels is useful in examining the benefits to society of future policy changes that may be introduced by eBay. Our study estimates that in the year 2003 alone, eBay online auctions contributed $1.5 billion in consumer surplus. We believe that these high and growing consumer welfare levels reveal why online auctions are increasingly popular.

REFERENCES

1. Adams, Christopher P. and Vogt, William (2004). "Digital Demand: Demand for digital cameras on eBay," available at www.ios.neu.edu/papers/s4g3.pdf

2. Adams, Christopher P. and Bivins, Laura L. (2003). "Focusing on Demand: Using eBay data to analyze the demand for telescopes," available at http://www.ftc.gov/be/workpapers/wp257.pdf

3. Ba, Sulin and Pavlou, Paul A. (2002). "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior," *MIS Quarterly*, 26(3): 269-289.

4. Bajari, Patrick and Hortaçsu, Ali (2003). "The Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights from eBay Auctions", *Rand Journal of Economics*, 3:2, pp. 329-355.

5. Bajari, Patrick and Hortaçsu, Ali (2004). "Economic Insights from Internet Auctions," forthcoming in *Journal of Economic Literature*.

6. Bapna, Ravi; Goes, Paulo; Gupta, Alok and Jin, Yiwei (2004). "User Heterogeneity and its Impact on Electronic Auction Market Design: An Empirical Exploration," *MIS Quarterly*, 28:1, pp. 21-43.

7. Bapna, Ravi (2003). "When snipers become predators: can mechanism design save online auctions?" *Communications of the ACM*, 46**:**12, pp. 152-158

8. Bapna, Ravi; Goes, Paulo; Gupta, Alok (2003a). "Replicating Online Yankee Auctions to Analyze Auctioneers' and Bidders' Strategies," *Information Systems Research*, 14:(3), 244-268.

9. Bapna, Ravi; Goes, Paulo; Gupta, Alok (2003b). "Analysis and Design of Business-to-Consumer Online Auctions" *Management Science*, 49:(1), 85-101

10. Brynjolfsson, Erik; Hu, Y. and Smith, Michael D. (2003). "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers," *Management Science,* 49:11, pp. 1580-1596.

11. Cohen, Marcel (2000), "Consumer Involvement – Driving up the Cost," *Consumer Policy Review*, 10(4): 122– 125.

12. David, Herbert Aron; Nagaraja, Haikady Navada (2003), *Order Statistics*, 3rd edition, John Wiley.

13. Diamond, Peter A. and Hausman, Jerry A. (1994). "Contingent Valuation: Is Some Number better than No Number?" *The Journal of Economic Perspectives*, Vol. 8, No. 4. (Autumn, 1994), pp. 45-64.

14. Ghose, A., M. Smith, R. Telang. (2004), "Internet Exchanges for Used Books: An Empirical Analysis of Welfare Implications," NYU Working paper available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=584401.

15. Giray, Tugba; Hasker, Kevin and Robin C. Sickles (2005), "Estimating Consumer Surplus in eBay Computer Monitor Auctions", paper presented at *FTC roundtable on The Economics of Internet Auctions*, Oct 2005.

16. Goolsbee, A., A. Petrin. (2001). The consumer gains from direct broadcast satellites and the competition with cable television. National Bureau of Economic Research Working paper W8317, Cambridge, MA.

17. Hasker, Kevin; Gonzalez, Raul and Robin C. Sickles (2005), "An Analysis of Strategic Behavior and Consumer Surplus in eBay Auctions," Rice University working paper, available at http://www.ruf.rice.edu/~rsickles/paper/auction.pdf

18. Hausman, Jerry A. (1981). "Exact Consumer's Surplus and Deadweight Loss," *The American Economic Review*, Vol. 71, No. 4. (Sep., 1981), pp. 662-676.

19. Jank, Wolfgang and Shmueli, Galit (2003). "Dynamic Profiling of Online Auctions using Curve Clustering", University of Maryland working paper, available at http://www.smith.umd.edu/ceme/statistics/AuctionProfiling.pdf

20. Katkar, Rama and Lucking-Reiley, David. (2000). "Public Versus Secret Reserve Prices in eBay Auctions: Results from a Pokémon Field Experiment," working paper, University of Arizona.

21. Klemperer, Paul (2002). "How (Not) to Run Auctions: the European 3G Telecom Auctions," forthcoming in *European Economic Review*.

22. Klemperer, Paul (1999), "Auction Theory: A Guide to Literature", *Journal of Economic Surveys*, 13(3), pp. 227-86.

23. Krishna, Vijay (2002), **Auction Theory**. Academic Press, San Diego, California..

24. Lambert, Diane (1992). "Zero-Inflated Poisson Regression, With an Application to Defects in Manufacturing", *Technometrics*, vol. 34, pp. 1-14.

25. List, John A. and Lucking-Reiley, David (2002) "Bidding Behavior and Decision Costs in Field Experiments" *Economic Inquiry*, vol. 40, no. 44, pp. 611-619

26. Lucking-Reiley, David (1999). "Using Field Experiments to Test Equivalence Between Auction Formats: Magic on the Internet*" American Economic Review*, vol. 89, no. 5, pp.1063-1080.

27. Lucking-Reiley, David (2000), "Auctions on the Internet: What's Being Auctioned, and How?" *Journal of Industrial Economics*, vol. 48, no. 3, pp. 227-252.

28. Lucking-Reiley, David; Bryan, Doug; Prasad, Naghi, and Reeves, Daniel (2000), ""Pennies from eBay: the Determinants of Price in Online Auctions," working paper, University of Arizona.

29. Marshall, A. (1936), *Principles of Economics*, Eight edition, London, pp. 124-133.

30. McAfee, R. Preston and McMillan, John (1987), "Auctions and Bidding," *Journal of Economic Literature,* Vol. 25, No. 2. pp. 699-738.

31. Milgrom, Paul R. and Weber, Robert J. (1982), "A Theory of Auctions and Competitive Bidding," *Economterica*, Vol. 50, No. 5, 1089-1122.

32. Myerson, Roger (1981) "Optimal Auction Design," *Mathematics of Operation Research*, 6, pp 58-73.

33. Nevo, A. (2001), New products, quality changes and welfare measures computed from estimated demand systems. Working paper W8425, National Bureau of Economic Research, Cambridge, MA.

34. Resnick, Paul and Zeckhauser, Richard (2001). "Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System," working paper, Harvard Kennedy School.

35. Roth, Al, and Ockenfels, A. (2002), "Last Minute Bidding and Rules for Ending Second-Price Auctions: Theory and Evidence from a Natural Experiment on the Internet," *American Economic Review*, 92 (4), pp 1093-1103.

36. Schindler, Julia. 2003. "Late Bidding on the Internet," mimeo, Vienna University of Economics and Business Administration.

37. Shmueli, Galit and Jank, Wolfgang (2004), "Visualizing Online Auctions," *Journal of Computational and Graphical Statistics*, 14 (2), pp. 299-319.

38. Shmueli, Galit, Russo Ralph P, and Jank, Wolfgang (2004), "Modeling Bid Arrivals in Online Auctions," University of Maryland Working paper. Available at http://www.smith.umd.edu/ceme/statistics/BidArrivals_Jan04.pdf

39. Smith, Vernon L., and Walker, James M. (1993), "Monetary Rewards and Decision Cost in Experimental Economics," *Economic Inquiry*, 31, pp. 245-261.

40. Eli M. Snir (2005) "Online Auctions Enabling the Secondary Computer Market," forthcoming in *Information Technology & Management*.

41. Song, Unjy. (2004) "Nonparametric Estimation of an eBay Auction Model with an Unknown Number of Bidders." available at: http://faculty.arts.ubc.ca/usong/eBay.pdf

42. Vickrey, William. (1961). "Counterspeculation, Auctions, and Competitive Sealed Tenders." *Journal of Finance*, 16:8-37.

43. Zeithammer, Robert (2005), "An Equilibrium Model of a Dynamic Auction Marketplace," available at

http://gsbwww.uchicago.edu/fac/robert.zeithammer/research/dynamic_auction_marketplace.pdf

ONLINE APPENDIX A

Validation Analysis

Our objective here is to test whether the sample of auctions obtained from Cniper.com is any different from a randomly selected sample of eBay auctions. To obtain the latter, we undertook a "title and description" advanced search of eBay auctions, in the three currencies, using a neutral phrase "May-13-04." The string representing this date returned the maximum number of listings in period of plus or minus 15 days, approximately 10,000[33]. Subsequently, after the last of these auctions closed, we obtained all the independent variable information by parsing the HTML pages of those auctions that had at least one bid submitted (1077 auctions[34]) to form the "validation data." The only data missing was the dependent variable (surplus), as eBay does not provide access to the winning bidder's bid. Subsequently, we compared the distribution of each of the independent variables in the validation data and the Cniper data to test for any significant difference.

These comparative studies are presented as a series of box plots and QQ plots for the numerical variables and bar- and pie-charts for the categorical variables. For all variables we found no significant difference between the Cniper data and the validation data, supporting the assumption that the Cniper data is no different than any other randomly drawn set of eBay data.

---

[33] It could have been the begin date, end date, modified by seller date or any other occurrence of the string "May-13-04" in the title or description of the auction listing.

[34] A large percentage of eBay auctions get no bids at all. Secondary eBay data sites such as www.andale.com and Hammertap.com's DeepAnalysis tool, give a range of so called "success rate" according to eBay category.

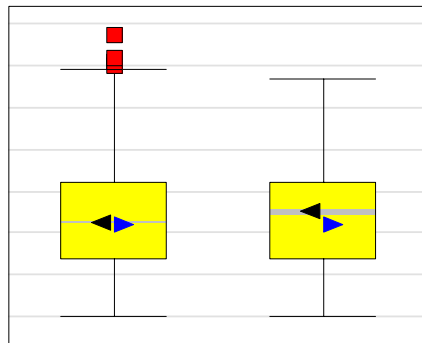**Box plots of Cniper variable vs. validation (eBay) variable.**
(black arrow = mean, blue arrow = median; left plot corresponds to
Cniper and the right to validation data; all variables are log transformed; ratings are also shifted
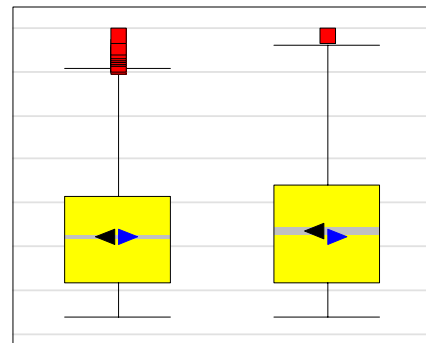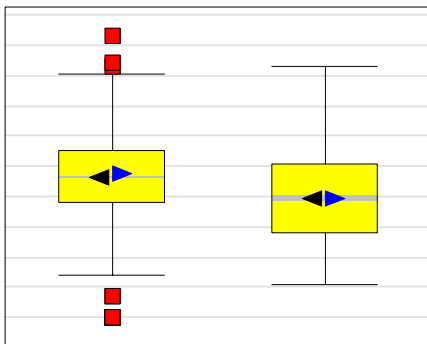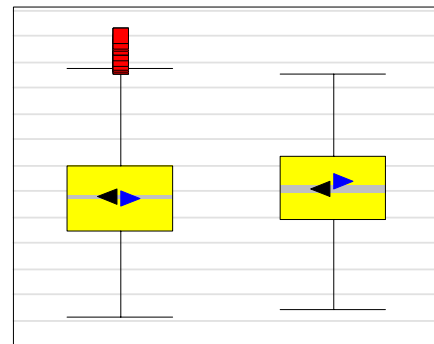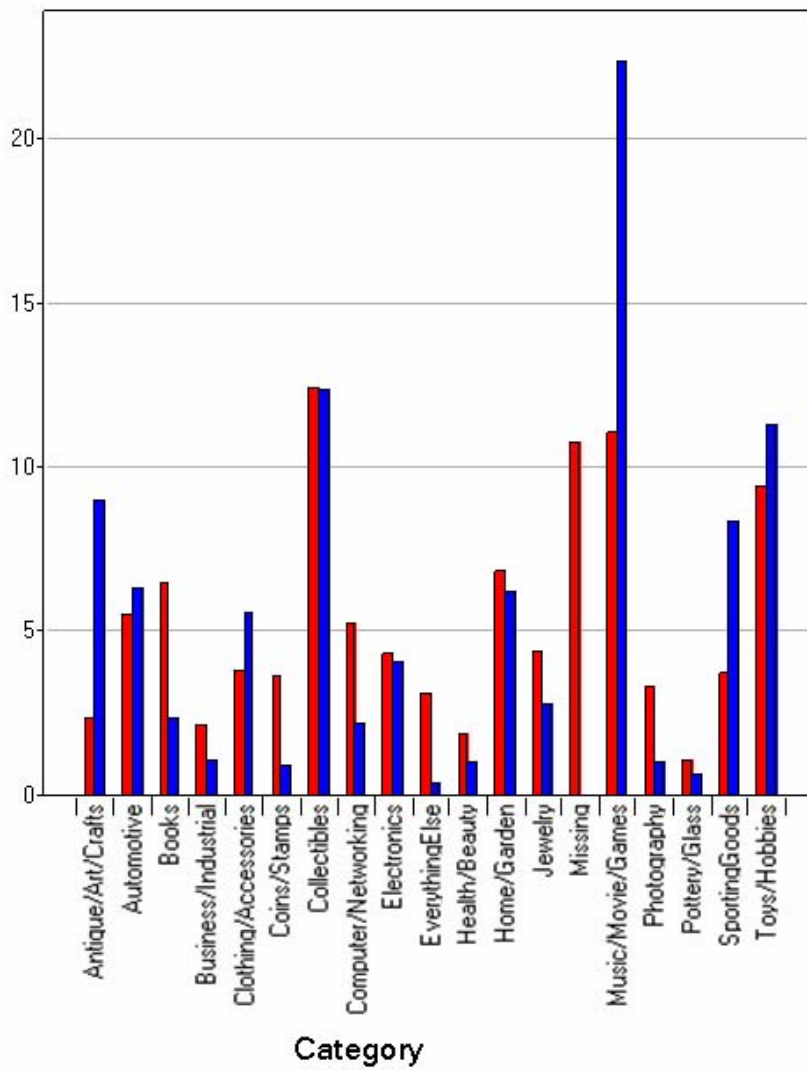by four to the right)
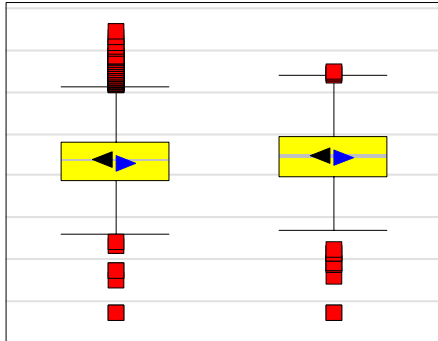
Price
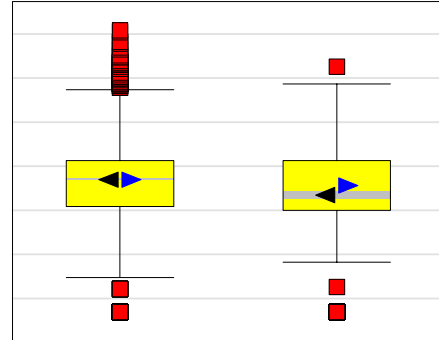
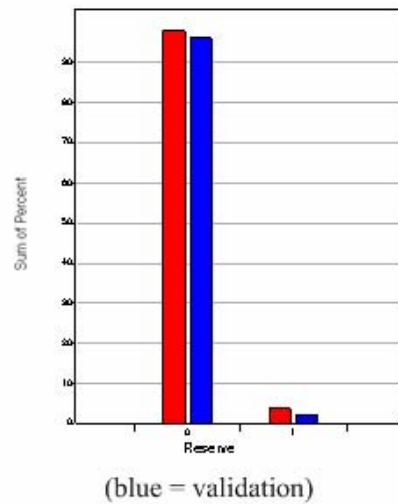
Opening Bid

Number of Bidders

Number of Bids
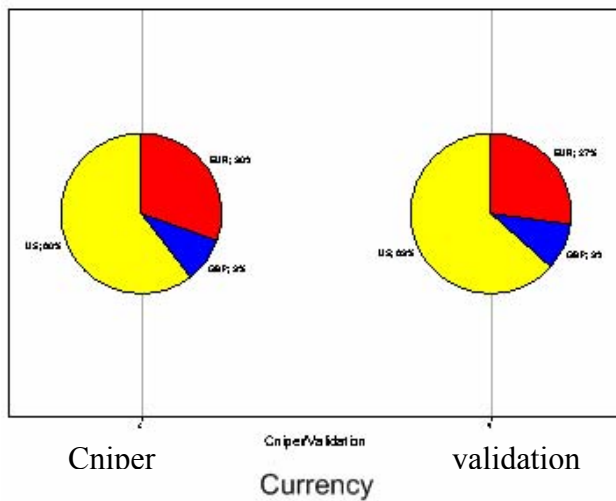
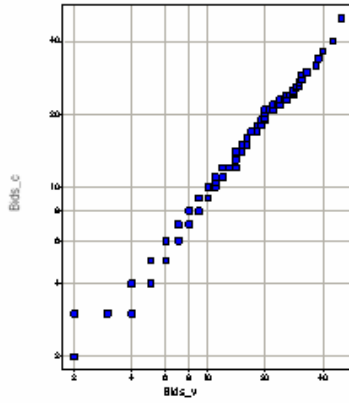

Winner rating (log +4)

Seller rating (log +4)

Categories (% auctions in each category). Blue is eBay validation data.

Cniper   validation

(blue = validation)

**QQ plots of Cniper variable vs. validation (eBay) variable.**

(Plots the Cniper percentile vs. its matching validation percentile (median vs. median, etc.) for each variable separately. Validation is on the horizontal axis. All variables are log transformed; ratings are also shifted by four to the right. A straight line of 45 degrees indicates that the distributions match. Note that most of the distributions closely match, and that discrepancies occur at the very top percentiles due to very right-skewed distributions.)
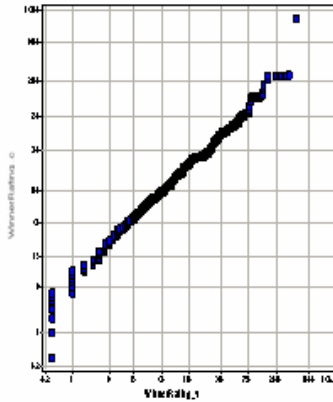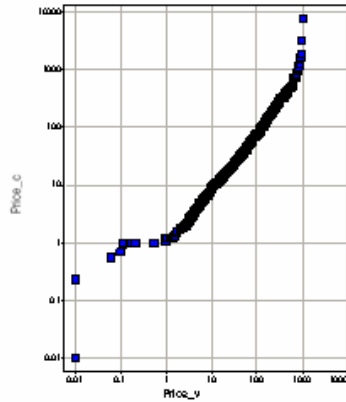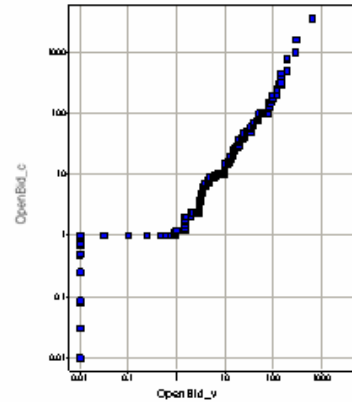

Number of Bids
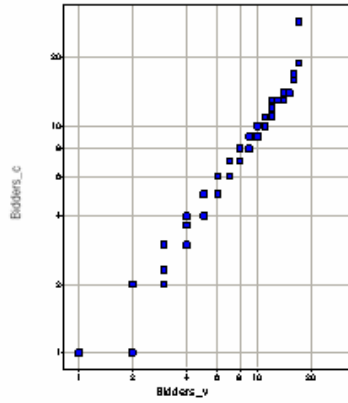
Number of Bidders
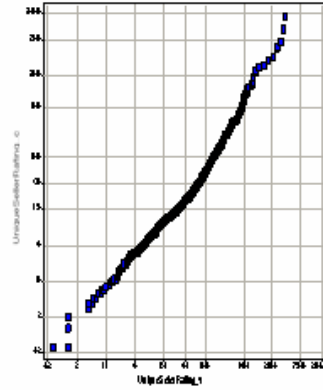
Seller rating

Winner Rating

Price

Opening Bid

ONLINE APPENDIX B – Two-Step Approach is Not as Powerful as the Unified Approach

Variable		Estimate	Std. Error	Pr(>|t|)
(Intercept)		9.3541	2.2461	0.0000
Categories				
	Antique/Art	0.6470	0.5203	0.2137
	Pottery/Glass	13.9756	347.5794	0.9679
	Collectibles	-0.0189	0.1900	0.9209
	Everything Else	0.7328	0.4657	0.1156
	Toys/Hobbies	0.2252	0.1885	0.2323
	Music/Movie/Games	0.2715	0.1749	0.1205
	Jewelry	-0.4403	0.3315	0.1840
	Automotive	0.2712	0.2333	0.2449
	Home/Garden	-0.3265	0.2063	0.1136
	Health/Beauty	0.3393	0.4748	0.4748
US DOLLARS		0.2778	0.1231	0.0240
NUM_DAYS		-0.6183	0.2857	0.0305
SNIPE_TIME		-0.7888	0.2405	0.0010
NUM_BIDDERS		-0.6323	0.2636	0.0164
PRICE		0.0627	0.1265	0.6202
S_RATING		-0.0502	0.0302	0.0963
W_RATING		0.0332	0.0421	0.4306
OPENING BID		-0.0298	0.0854	0.7272
OPENING BID x PRICE		-0.0072	0.0191	0.7041
PRICE x NUM_BIDDERS		0.0616	0.0715	0.3889
NUM_DAYS x SNIPE_TIME		0.0734	0.0321	0.0222

Table B-1: Parameter estimates from the logistic regression model on positive vs. zero surplus

Variable		Estimate	Std. Error	Pr(>|t|)
(Intercept)		2.7085	0.8124	0.0009
Categories				
	Antique/Art	0.4936	0.1544	0.0014
	Pottery/Glass	0.2425	0.2225	0.2757
	Collectibles	0.6866	0.0791	0.0000
	Everything Else	0.4412	0.1382	0.0014
	Toys/Hobbies	0.2101	0.0837	0.0121
	Music/Movie/Games	0.4149	0.0837	0.0000
	Jewelry	-0.3531	0.1191	0.0030
	Automotive	-0.3298	0.1076	0.0022
	Home/Garden	-0.3534	0.0973	0.0003
	Health/Beauty	-0.5300	0.1749	0.0025
US DOLLARS		0.3031	0.0535	0.0000
NUM_DAYS		-0.2616	0.1052	0.0129
SNIPE_TIME		-0.3599	0.0870	0.0000
NUM_BIDDERS		-0.2990	0.1100	0.0066
PRICE		0.5882	0.0501	0.0000
S_RATING		-0.0614	0.0131	0.0000
W_RATING		0.0277	0.0177	0.1169
OPENING BID		-0.1360	0.0354	0.0001
OPENING BID x PRICE		0.0272	0.0077	0.0004
PRICE x NUM_BIDDERS		0.0420	0.0288	0.1453
NUM_DAYS x SNIPE_TIME		0.0371	0.0121	0.0021

Table B-2: Parameter estimates from a regression model on positive surplus only.

Table B-1 and B-2 show the results from a two-step estimation approach. First we estimate the occurrence of positive surplus using a logistic regression model on an indicator variable that equals one if the surplus is positive and zero otherwise. Table 1 shows the parameter estimates. We can see that the results are very similar to our estimation approach. Almost all parameters have the same sign. Notice that only the parameters for Automotive, Health and Beauty and the interaction between the opening bid and price have signs opposite to our model. Notice also though that all of these three variables are highly insignificant! In fact, the logistic regression model features many more insignificant variables compared to our approach. The reason for this is the loss in power when predicting a dichotomous variable rather than a continuous factor.

Table B-2 shows the result form the second step. Again, all parameter estimates have the same sign as in our model. However, notice the loss in significance for many of the parameters. In particular, Pottery/Glass and the winner's rating are now only borderline significant. Overall, the two-step approach leads to a much less powerful model than the unified approach proposed in our manuscript.