Tests of the sealed-bid abstraction in online auctions

Robert Zeithammer
UCLA

Christopher Adams
Federal Trade Commission

August 2, 2008

Abstract

This paper presents five empirical tests of the popular modeling abstraction that bidders in ascending online auctions bid "as if" they were in a sealed bid auction. The tests rely on observations of the magnitudes and timings of top two proxy bids, with the different tests stemming from different regularity assumptions about the underlying joint distribution of signals and timings. We apply the tests to data from three eBay markets - MP3 players, DVDs and used cars - and we reject the sealed-bid abstraction in all three datasets. This consistent rejection casts doubt on several existing theories of online-auction behavior. Moreover, we reject the sealed-bid abstraction even in carefully selected subsets of the data that one might consider more likely to conform to sealed bidding. Given these findings, demand-estimation using eBay data is more difficult than previously thought. In particular, our results suggest that the empirical strategies based on multiple order-statistics of the bidding distribution will not work.

We thank Raphael Thomadsen, Jeremy Fox, Ali Hortacsu, Harry Paarsch, and seminar participants at Chicago GSB for comments and suggestions. We also thank eBay for providing the data, and Uri Simonsohn for cleaning up and sharing the movie dataset. The views expressed do not necessarily reflect the views of the FTC or any individual Commissioner. All errors are our own. Contact: Robert Zeithammer, UCLA Anderson School of Management, 110 Westwood Plaza, Los Angeles, CA 90095. Email: rzeitham@ucla.edu.

1. Introduction: the sealed-bid abstraction

The rules of most online auctions, such as eBay, resemble an ascending English auction because the bidders can revise their bids upward whenever they are outbid. However, all major auction sites provide proxy-bidding software designed to save the bidder's effort by automatically bidding on his behalf up to a secret maximum. Because of this proxy system, the economically relevant endgame of each auction resembles a simple second-price sealed-bid auction, and so it seems reasonable to model online auctions as second-price sealed-bid auctions. We propose and apply a series of empirical tests of this sealed-bid abstraction in models of online-auction bidding. Testing this particular abstraction is important for two reasons: First, our current theoretical understanding of online auctions is dominated by models that involve the abstraction as their main simplification. A rejection of the abstraction thus focuses future theorists on more complicated theories that do not abstract from within-auction dynamics. Second, the state-of-art method for nonparametric identification of demand in online auctions depends crucially on the abstraction. Measuring the structural primitives of demand is one of the central goals quantitative analysis in Marketing and Economics, and online auctions potentially contain a wealth of demand information for both researchers and managers (Chakravarti et al 2002). A rejection of the abstraction thus focuses future demand analysts on robust econometric methods that can measure demand even when bidding is not actually sealed. Before summarizing our results, we discuss the two reasons for testing the sealed-bid abstraction in more detail.

There are at least six existing theories of online bidding that involve the sealed-bid abstraction. Most of the theories that take bid-timing into account at all predict bidding at the last moment of the auction. Placing a proxy-bid (equal to an equilibrium bid in the corresponding second-price sealed-bid auction) at the last moment can be a symmetric equilibrium strategy for several reasons: First, this strategy can sometimes allow tacit collusion against the seller (Roth & Ockenfels 2002). Second, it protects private information in a common-value setting (Bajari & Hortacsu 2003). Third, it avoids bidding wars with an irrational fringe of "incremental" bidders (Ariely, Ockenfels & Roth 2005). Fourth, it maximizes the information about upcoming options to buy substitute products (Zeithammer 2006). Several theories simply abstract away from the timing of bids under the assumption of independent private values. For example, Yao and Mela (2008) appeal to the equivalence between the truth-revealing symmetric equilibrium of a second-price sealed-bid auction and an equilibrium of an ascending English auction. Alternatively, Song

(2004) assumes that bidders do not have the luxury of waiting until the last moment but instead each have a different last opportunity to bid. Placing a proxy-bid equal to one's valuation[1] at one's last opportunity to bid can be an equilibrium strategy of the online auction game (Song 2004). Please see the excellent survey by Bajari & Hortacsu (2004) for an in-depth exposition of most of the aforementioned theories. Given the theoretical agreement that the sealed-bid abstraction should describe online bidding, testing it empirically sheds light on how well extant theory captures actual behavior. A rejection of the sealed-bid abstraction leaves several models that do not depend on it. For example, Bradlow and Park (2007) propose a latent stochastic process for bid-increments and bidder-arrivals, with observed bids evolving over time analogously to a record-breaking process. Existing equilibrium models that do not predict sealed-like bidding usually rely on the observation that there are multiple concurrent auctions on eBay: Peters and Severinov (2006) obtain non-sealed bidding in an equilibrium of a model with sequential arrivals of bidders to many simultaneous auctions. In contrast to Song (2004) who studied a single auction, Peters and Severinov show that "the presence of multiple auctions implies that it is not a dominant strategy and not even a sequentially rational strategy in a perfect Bayesian equilibrium for buyers to bid their true valuations when they start bidding." (p. 225). Nekipelov (2007) also focuses on multiple concurrent auctions, and shows that there is an incentive to bid early in order to deter entry by rival bidders. Dynamics that do not satisfy the sealed-bid abstraction emerge in Nekipelov's model.

The sealed-bid abstraction is not only consistent with extant theory, but it is also critical for nonparametric identification of demand from online auction data. In standard auction-models, demand is nonparametrically identified by the closing price and the number of bidders (Athey & Haile 2002, 2005). When the number of bidders is unobservable, Song (2004) argues that identification is still possible if two order-statistics of the bid-distribution are observed. This identification strategy relies heavily on the sealed-bid abstraction, so empirical researchers need a test to determine whether the abstraction holds in their particular datasets. A rejection of the sealed-bid abstraction points to methods that do not rely on it, and instead model the latent number of bidders. For example, Chan et al (2007) extend the bounds approach of Haile and Tamer (2003) to account for latent bidders as observed bidders in concurrent auctions.

[1] Throughout this paper, "consumer's valuation" means the consumer's maximum willingness to pay at the moment, i.e. the dollar utility of the good *net* of all other opportunities to buy a substitute good elsewhere. Specifically, valuation is not the consumer's intrinsic value of the product (see also Bajari and Hortacsu 2003, Chan et al 2007).

Alternatively, Adams (2007) relies on an exogenous proxy for the number of bidders. Having discussed the importance of the sealed-bid abstraction, we now turn to our tests.

The proposed tests use data on magnitudes and timing of the top two bids in each auction, where "bid" stands for the secret maximum proxy-bid. Under the null hypothesis that the sealed-bid abstraction holds, the following five things should be true: T1) the top bid should be equally likely to arrive before or after the second-highest bid, T2) the top two bids should not be exactly one bid-increment apart too often, T3) the difference between the top two bids should not be a function of which was placed first, T4) the difference between the top two bids should not be a function of time remaining in the auction, and T5) the conditional distribution of the top bid given the second-highest bid should have the same right tail for any particular value of the second-highest bid. In this paper, we specify the different assumptions under which each of these five tests works, and we propose a way to non-parametrically operationalize the novel test T5 which does not depend on timing or bid-increment data. We then apply all tests to three different datasets from eBay (MP3 players, movies on DVD, and used cars).

Our data is provided directly by eBay, and so we observe the proxy bid of the winner. This information is critical to our tests, but it is not available from the eBay website. Therefore, we report on empirical regularities of uniquely complete datasets.[2] Theoretically, there is nothing special about the top two bids in an auction – most of the above tests would be valid with any other pair of order-statistics of the bidding distribution. Practically, however, the top two order-statistics are the only reliably observable ones because of truncation issues: On eBay, a bid can be submitted only if it exceeds the highest bid at the moment, so eBay data contains relatively more high bids and relatively fewer low bids than a random sample from the underlying population distribution of bids. Although many latent bids may thus be truncated, two bids in every auction are always recorded: the highest and the second-highest bid in each auction. Because eBay does not allow a bidder to outbid himself, the two order-statistics correspond to bids submitted by different people. Therefore, the first- and second-order statistics $b_1 \geq b_2$ of the population distribution of bids are always observed, along with the times at which they were placed (t_1, t_2), where t_i is the time at which bid b_i was submitted. Conversely, the lower order-

[2] Obtaining data directly from eBay is not the only way to observe the proxy bid of the winner. Bapna, Jank & Shmueli (2008) observe it by running their own sniping agent. Alternatively, researchers willing to win auctions with high last-second bids will observe the proxy-bid of the highest bidder from the price they pay.

statistics of the observed bids do not correspond to the order-statistics of the latent bids (see Song 2004 for an in-depth discussion). Given the truncation issue, we take the (t_1, t_2, b_1, b_2) observations as a starting point, and develop a series of tests of the sealed-bid abstraction.

Another reason to focus on the top two bids is the obvious alternative "ascending" model of bidding, in which each bidder manually raises his proxy bid up to his maximum. The top two bids are special in that ascending bidding by all participants predicts sharply different outcomes of T1-T5 from sealed bidding. Specifically, when bidders bid in an ascending fashion, the following should be true: in T1, the top bid should always arrive after the second-highest bid, so the test T3 degenerates to a single cell. In T2, the top two bids should always be exactly one bid-increment apart, so the conditional distribution of the top bid given the second-highest bid (considered in T4 and T5) should degenerate to a mass-point one increment above the second-highest bid. The minimum bidding increment *inc* is variable but common knowledge in all standard online auctions, so it can be considered as part of the data. Therefore, some of our tests rely on observations of $(t_1, t_2, b_1, b_2, inc)$.

Taken together, our application of all five tests rejects the sealed-bid abstraction as a general property of bidding in eBay auctions. Three tests reject consistently, with surprising empirical regularity of the test-statistics across the three diverse datasets: First (T1), b_1 is placed after b_2 in about two thirds of the auctions. Second (T2), about fifteen percent of the auctions end with the two bids exactly one increment apart (compared to only four percent that end in an exact tie). Finally (T3), the exactly-increment-apart outcome is about three times more likely when b_1 comes after b_2 compared to the reverse order. The easiest way to explain these findings is that significant proportion (about 30 percent) of the auctions within each dataset are better described by ascending bidding. While sealed bidding is thus not a general property of all auctions, it may still describe a subset of auctions large-enough for demand-estimation ala Song. We explore this possibility with two "plausibly sealed" candidate subsets: the "OverInc" auctions in which b_1 exceeds b_2 by more than an increment, and the "HighFirst" auctions in which b_1 is placed before b_2. Interestingly but sadly for nonparametric demand identification, we can still convincingly reject the sealed-bid abstraction in both candidate subsets based on the joint distribution of bids and timing. However, the selection rules of these subsets exclude several of the tests. For example, test T2 does not work in the OverInc data while test T1 does not work in the HighFirst data. Therefore, we need the more fine-grained test T5 which does not rely on timing data and is applicable across all data-subsets.

This last test (T5) is a nontrivial contribution to the empirical auction literature. Instead of relying on timing or increment data, it considers only the joint distribution of (b_1, b_2) and asks what should be true about the residuals of the demand-identifying regression of b_1 truncated at b_2. The sealed-bid abstraction implies a powerful regularity: the right tail of the distribution of the residual of b_1 conditional on the residual of b_2 should have the same shape regardless of the particular value of the residual of b_2. We operationalize this idea using a new nonparametric test, and apply it to both OverInc and HighFirst subsets of auctions. The results of this test are more nuanced than those of the first four tests: the sealed-bid abstraction is rejected more often and more strongly in the OverInc data than in the HighFirst subset of the OverInc data. Moreover, differences across the datasets also emerge – the bidding behavior in car auctions seems to be qualitatively different from the bidding behavior in MP3-player and DVD-movie auctions.

The paper is organized as follows: the next section introduces the test and the assumptions they are based on. Section 3 then describes our data and applies the tests. Section 4 considers robustness of our results to relaxation of the assumptions, and Section 5 discusses the implications of our findings for demand-estimation with eBay data. Section 6 then concludes by summarizing our results, and outlining how the empirical regularities we document constrain theories of online auction behavior.

2. Tests of the sealed-bid abstraction in online auctions

2.1 Definitions

Our tests are geared towards testing whether the economically important bids in eBay auctions behave analogously to bids in a second-price sealed-bid auction, but the tests' applicability extends to other auctions and other sealed-bid scenarios. To facilitate the widest possible scope of application, we define our primitives in maximum generality.

Our concept of an *online auction* is very broad, and encompasses any auction that receives bids over time, with each bid associated with a unique time-stamp. For the purposes of this paper, every current Internet auction is thus an online auction, but so is every other auction that receives mail-in or call-in bids. The particulars of feedback to bidders during the auction are not important, and they can range from none (as in a government sealed-bid auction with mail-in bids) to a full record of successful bids to date (as in an eBay auction).

The main focus of this paper is sealed bidding as a model of online auctions. A model of an online auction involves a *sealed-bid abstraction* whenever the bidders bid as if they were in a standard sealed-bid auction. Mathematically, each bidder in a sealed-bid auction receives a private scalar signal x, and bids according to a strictly-increasing function $\beta(x): x \rightarrow bid$ that depends only on x. For example, Bajari & Hortacsu (2003) assume that x is a private signal about the common value of the auctioned good (a collectible coin). In such a common-value environment, there is a symmetric equilibrium, in which all bidders bid at the last moment as if they were in a second-price sealed-bid auction with common values. The equilibrium bid-function is increasing and mitigates winner's curse by $\beta(x) < x$. Similarly, bidders in Song's (2004) model receive independent signals about their private valuations, and each bidder has an exogenous last opportunity to bid. In such a private-value model, the online auction becomes strategically equivalent to a second-price sealed-bid auction, and so each bidder has a dominant strategy to bid x at his last opportunity to bid. Note that in both examples, the second-price nature of the sealed-bid auction arises from eBay's proxy-bidding agent. While the previous two examples involve equilibrium models of bidding, our tests do not depend on the equilibrium assumption - $\beta(x)$ could be any ad-hoc behavioral regularity. For example, bidders may not actually be strategic but they may follow eBay's instructions that direct them to submit their maximum willingness to pay as their proxy bids. Having defined the two key primitives of our theory, we now turn to the properties of auction models from which our tests arise.

2.2 Model properties

We will present a series of tests, and different tests will rely on different properties of the auction model. Therefore, even if one or two of the following properties do not hold, some of the proposed tests will still work. The full list of properties we will be using is:

- A1 (timing independent of signals): t_i is independent from x_i for every bidder *i*.
- A2 (continuity): signals x_{ij} of bidder *i* in auction *j* are drawn from a continuous distribution
- A3 (conditional iid): conditional on auction-specific observables Z_j, signals x_{ij} are distributed *iid* across auctions *j* and bidders *i*.
- A4 (increment): the auction is actually an ascending auction with a minimum bid-increment *inc* which is observed by the econometrician.

The first property (A1) says that there is no link between the magnitude of private signals and the timing of the bids associated with those private signals. As discussed in the introduction, this property does not hold in the models of Bradlow and Park (2007), Peters and Severinov (2006), and Nekipelov (2008). The second property (A2) says that the signals are drawn from a continuous distribution which may vary from auction to auction and may involve arbitrary correlations across bidders and/or auctions. While many existing auction models have this property, it is an unrealistic assumption about the auction environment when some bidders value goods in whole dollar amounts. In the theory literature, the model of Peters and Severinov (2007) actually relies on discreteness of the distribution for the equilibrium to exist. Section 4 will generalize A2 to allow mass-points in the distribution and discuss how such a generalization would impact testing. A popular property that allows pooling of data across auctions is that the signals are distributed *iid* across auctions and bidders. This “iid” property is property A3, except that A3 conditions on auction-level observables. Finally, property A4 is an assumption about the rules of the auction that generates the data.

2.3 Tests

Suppose we have the data on top two bids in an auction $b_1 \geq b_2$, their timing $t_i = time(b_i)$, and the increment *inc*. In the eBay setting, (b_1, b_2) are the top two proxy-bids submitted in the auction and *inc* is eBay's minimum increment[3]. Given data on $(t_1, t_2, b_1, b_2, inc)$, the simplest test is based on timing alone: as long as timing of bids is independent of private signals (A1), the sealed-bid abstraction predicts that neither of the bids should be more likely to appear first:

$$(\text{A1}) \Rightarrow (\text{T1}) : \Pr(t_1 > t_2) = \frac{1}{2}$$

In the alternative ascending-bidding model, $\Pr(t_1 > t_2) = 1$, so any empirical probabilities $\Pr(t_1 > t_2) > \frac{1}{2}$ suggest a departure towards ascending bidding. To operationalize this test, we compute the empirical probabilities with their associated standard errors.

In online ascending auctions with a minimum increment (A4), another simple test uses only data on bids and the continuity property (A2):

$$(\text{A2,A4}) \Rightarrow (\text{T2}) : \Pr(\Delta b = inc) = 0 .$$

In contrast, the ascending model predicts that $\Pr(\Delta b = inc) = 1$, so T2 failing should be a strong indication of ascending behavior. To operationalize T2, we again compute the empirical probability with its standard error. As pointed out in the previous section, the continuity assumption may be too strong because the actual distribution may have mass-points at round numbers: bidders may only think of their valuations in whole-dollar amounts, or their valuations may congregate on the price available in an outside spot-market like Amazon. With mass-points, $\Pr(\Delta b = inc) > 0$, and test T2 becomes weaker. Please see Section 4 for a generalization of the T2's idea to a distribution with mass-points.

Combining the bid-data with the timing-data allows for more detailed testing based on the independence assumption (A1). By definition of independence between timing and bidding, A1 implies that Δb should be completely invariant to the timing of the bids. Two fruitful tests

---

[3] eBay's increment is a function of b_2 and varies from 5¢ to \$100 dollars as b_2 increases from \$1 to \$5000.

arise, again motivated by the ascending-behavior alternative: First, the sealed-bid abstraction predicts that the difference between the top two bids should not depend on which bid came first:

$$(\mathrm{A1}) \Rightarrow (\mathrm{T3}): cdf\left(\Delta b \mid t_1 > t_2\right) = cdf\left(\Delta b \mid t_1 < t_2\right).$$

If some auctions behave more like ascending auctions with $\left(t_1 > t_2\right)$ and others more like sealed-bid auctions, then $\left(\Delta b \mid t_1 > t_2\right)$ will be smaller and related to the minimum bid-increment while $\left(\Delta b \mid t_1 < t_2\right)$ will be "more continuous". To operationalize this test-idea, we use the nonparametric Wilcoxon-Mann-Whitney rank-test (hereafter WMW) and compare the subsample $\left(\Delta b \mid t_1 > t_2\right)$ to the subsample $\left(\Delta b \mid t_1 < t_2\right)$.[4] Analogously with conditioning on the relative timing of the top two bids, conditioning on time remaining in the auction should also leave Δb unchanged. When $\bar{t}$ is the ending time of the auction, define the time left as $t^R \equiv \bar{t} - \min\left(t_1, t_2\right)$, and the sealed-bid abstraction implies:

$$(\mathrm{A1}) \Rightarrow (\mathrm{T4}): cdf\left[\Delta b \mid t^R > median\left(t^R\right)\right] = cdf\left[\Delta b \mid t^R < median\left(t^R\right)\right].$$

To operationalize this test, we use the same mixture of non-parametric and parametric statistical methods as in T3.

Bringing in the increment information allows us to zoom in on the value of $\Delta b = inc$ to conduct special cases of T3 and T4 with a sharp alternative hypothesis under the ascending model. Specifically, we can test whether

$$(\mathrm{A1,A4}) \Rightarrow (\mathrm{T3'}): \Pr\left(\Delta b = inc \mid t_1 > t_2\right) = \Pr\left(\Delta b = inc \mid t_1 < t_2\right)$$
$$(\mathrm{A1,A4}) \Rightarrow (\mathrm{T4'}): \Pr\left[\Delta b = inc \mid t^R > median\left(t^R\right)\right] = \Pr\left[\Delta b = inc \mid t^R < median\left(t^R\right)\right]$$

Note that all of the above tests do not make any assumptions about the correlation of signals, either within or across auctions. Focus on a single auction *j*, and suppose (as in A3) that the signals are only correlated within the auction because of some common shock Z_j which is observable to both the bidders and the econometrician (on eBay, auctions differ in ending time, characteristics of the good sold, and characteristics of the seller). As long as Z_j is the only source

[4] In parallel, we also computed the standard t-test of the $E\left(\log \Delta b \mid t_1 > t_2\right) = E\left(\log \Delta b \mid t_1 < t_2\right)$ hypothesis, implicitly assuming log-normality of Δb. This parametric test almost perfectly agrees with the WMW test.

of dependence, the sealed-bid residuals ε_{ij} after conditioning the bids on Z_j are distributed *iid* according to some distribution F_j. The second part of A3 then allows pooling of data across auctions by assuming that the distribution of residuals is the same in all auctions in the sample:

$$\left(\text{A3 within auction}\right) \Rightarrow \varepsilon_{ij} \equiv \beta\left(x_{ij}\right) - E_i\left[\beta\left(x_{ij}\right) | Z_j\right] \overset{iid}{\sim} F_j$$
$$\left(\text{A3 across auctions}\right) \Rightarrow \forall k, j : F_j = F_j \equiv F$$

The residual ε_{ij} is the component of the bid $\beta\left(x_{ij}\right)$ due to private information of bidder *i* in auction *j*. *F* is the distribution of "demand" that demand-analysts seek to recover from bidding data: when the auction sells private-value goods, *F* is the distribution of private valuations in the population. In a common-value setting, *F* is the distribution of bids that all correct for the winner's curse by bidding below the private signal of value, so *F* can be used together with $\beta^{-1}\left(\cdot\right)$ to recover the population distribution of private signals. The above *iid* assumption is at the heart of Song's nonparametric strategy for recovering *F* (Song 2004). Specifically, Song shows that knowing any two order-statistics of the bidding distribution and their order is sufficient for identification of *F* even when the analyst does not know the number of bidders in each auction. When the two observed order-statistics of bids are $\left(b_1, b_2\right)$ - as in our situation - the identification of *F* is particularly simple: $\Pr\left(\varepsilon_1 < z | \varepsilon_2 = w\right) \sim \frac{F\left(x\right) - F\left(w\right)}{1 - F\left(w\right)}$, where the highest residual ε_1 is the residual of b_1 and ε_2 is the residual of b_2. In particular, the conditional distribution of ε_1 given ε_2 fixed at *w* is just the right tail of *F* truncated at *w*: $pdf\left(\varepsilon_1 | \varepsilon_2 = w\right) = \frac{f\left(\varepsilon_1\right)}{1 - F\left(w\right)}$. Independence across bidders within the auction is a key assumption for this result to go through, *iid* across auctions is only useful to pool data across auctions.[5] This regularity implies an entire range of tests based on the fact that the shape of the conditional distribution of $\left(\varepsilon_1 | \varepsilon_2\right)$ does not depend on the particular value of ε_2:

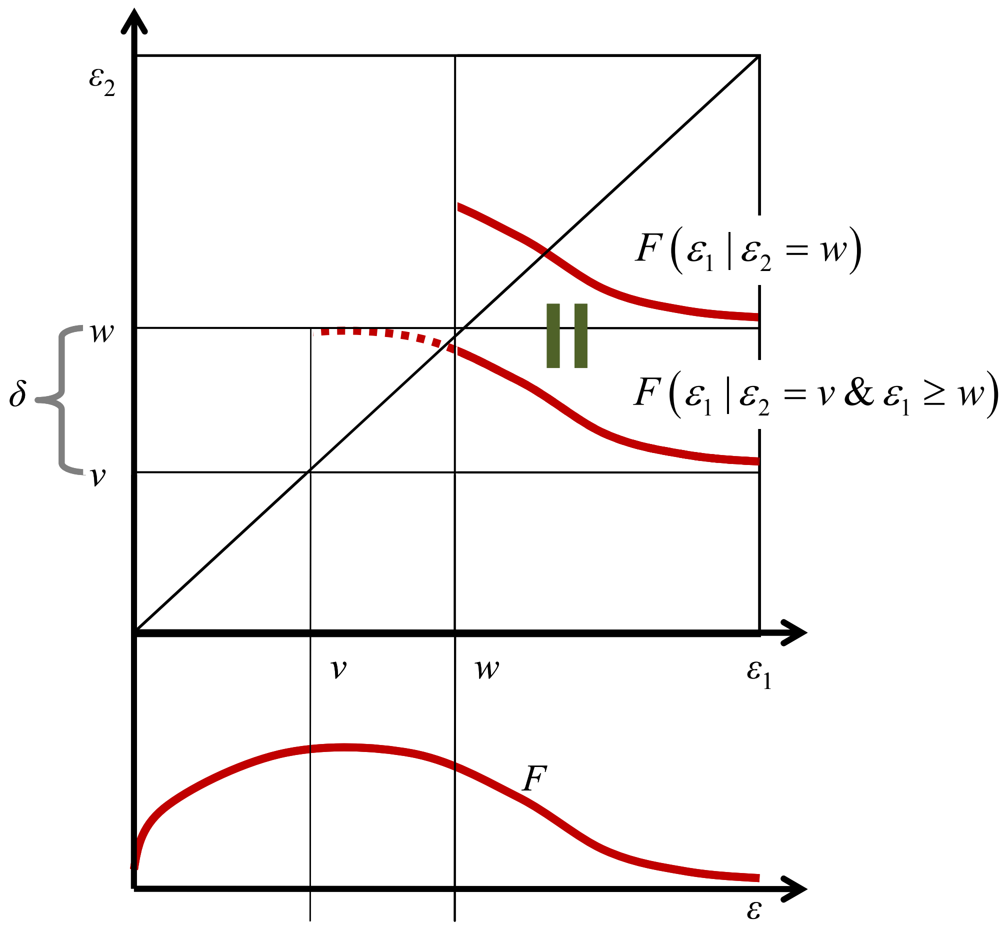
---

[5] Independence implies the joint distribution $pdf\left(\varepsilon_1, \varepsilon_2\right) = N\left(N-1\right) f\left(\varepsilon_1\right) f\left(\varepsilon_2\right) F\left(\varepsilon_2\right)^{N-2}$, i.e. the probability of the event that one draw is ε_1, another draw is ε_2 and all other draws are below ε_2. The resulting multiplicative form of the density separates the two order statistics. Thus, when ε_2 is a constant *w*, $pdf\left(\varepsilon_1, \varepsilon_2 | \varepsilon_2 = w\right) \propto f\left(\varepsilon_1\right)$.

$$(\text{A3}) \Rightarrow (\text{T5}): \forall w \,\&\, \delta > 0: cdf\left(\varepsilon_1 \mid \varepsilon_2 = w\right) = cdf\left(\varepsilon_1 \mid \varepsilon_2 = w - \delta \,\&\, \varepsilon_1 \geq w\right)$$

Figure 1 illustrates the resulting "tail-comparison test" for a particular value of *w* and *δ*, with the equal sign corresponding to the equal sign in the above equation. Table A1 in the Appendix illustrates what the conditional distributions of $\left(b_1 \mid b_2\right)$ look like in practice for a specific value of *w* (in raw bid-data for the most popular movie title) , i.e. without conditioning on auction-level observables Z_j).

Figure 1: Illustration of the tail-comparison test



Note to Figure 1: The lower part of the graph shows the population distribution F of the private bid-residual ε. The upper part of the graph shows the distribution of the ε_1 given ε_2 for two values of ε_2 : $w=v+\delta > v$. Note how the non-dashed tails of the two conditional distributions are the same as the tail of F truncated below at *w,* and so equal to each other.

Theoretically, every possible w could produce a separate test like the one in Figure 1, but it is not clear how one could use all the data jointly. To achieve this, we develop a new nonparametric statistical test based on all pairs of auctions: Consider all pairs of auctions j and k such that $\min\left(\varepsilon_{1,j},\varepsilon_{1,k}\right)>\max\left(\varepsilon_{2,j},\varepsilon_{2,k}\right)$, and call such pairs the "feasible pairs". The prediction of T5 is that in all the feasible pairs, the auction with a relatively higher price is <u>not</u> more likely to also have a higher top bid:

$$\text{T5 statistic} = \Pr\left[\varepsilon_{1,j}>\varepsilon_{1,k} \mid \varepsilon_{2,j}>\varepsilon_{2,k} \;\&\; \min\left(\varepsilon_{1,j},\varepsilon_{1,k}\right)>\max\left(\varepsilon_{2,j},\varepsilon_{2,k}\right)\right]=\frac{1}{2}$$

To test this prediction using the entire data-sample, we simply compute an empirical estimate of the above probability across all feasible pairs, and compare it to ½. To gauge statistical significance, we bootstrap the test-statistic using 10,000 re-samples with replacement from the original dataset (Efron and Tibshirani 1994). Please see the Appendix for details.

We have computed the T5 test-statistic under the null (*iid*) hypothesis while assuming different parametric data-generating distributions F, and we found that the distribution of the test-statistic under the null hypothesis seems to be invariant to the distribution that generates the bids. Therefore, we conjecture that T5 is a non-parametric test in that its distribution under the null does not depend on the distribution of bids in the population. Please see the Appendix for additional properties of the statistic. We believe that this test-statistic is new, and potentially useful outside of auction research. Whenever an analyst has data that naturally groups into pairs of order-statistics, the test-statistic can be used to test whether the data is iid across and within pairs without making any specific distributional assumptions. For example, suppose I have data on the heights of the two tallest children in many families that all have at least two children (but can differ in the number of children). I can use T5 to check whether child heights are iid in the entire population (and I will probably reject this hypothesis because height is partly genetic).

The null-hypothesis prediction of T5=½ is completely non-parametric, and it does not depend on assumptions A1 or A2. However, large samples of exactly identical auctions are rare, so non-parametric conditioning on Z_j is not practical. It is thus necessary to operationalize T5 as a semi-parametric test with the conditioning on Z_j achieved though a parametric model. We assume that bids are distributed log-normally and the effect of Z_j on bids is additively separable from the private bidder-specific residuals: $\log\beta\left(x_{i,j}\right)=\alpha Z_j+\varepsilon_{i,j}$ and $\varepsilon_{i,j}\sim N\left(0,\sigma^2\right)$

Under this specification, a Normal regression of $\log b_{1,j}$ on Z_j truncated at $\log b_{2,j}$ recovers α, and the residuals $\left(\hat{\varepsilon}_{1,j}, \hat{\varepsilon}_{2,j}\right)$ can be used to compute the test-statistic T5. One practical problem with a truncated Normal regression is lack of convergence in likelihood-maximization when the model is severely-enough misspecified. Whenever we cannot achieve convergence, we resort to simply regressing $\log b_{2,j}$ on Z_j and applying the test to the resulting OLS residuals. To further reduce the potential confounding impact of unobserved heterogeneity, we also perform the test separately for every product, i.e. separately for different movie-titles and MP3-player models.

This concludes the theoretical development of econometric tests, Table 1 illustrates how the different tests depend on different subsets of properties (assumptions).

Table 1: Which tests rely in which properties

	Properties (assumptions of tests)			
Tests:	*A1: timing inde-pendent*	*A2: signals conti-nuous*	*A3: signals cond. iid*	*A4: inc known*
$\text{T1}: \Pr\left(t_1 > t_2\right) = \frac{1}{2}$	✔			
$\text{T2}: \Pr\left(\Delta b = inc\right) = 0$		✔		✔
$\text{T3}: cdf\left(\Delta b \mid t_1 > t_2\right) = cdf\left(\Delta b \mid t_1 < t_2\right)$	✔			
$\text{T3'}: \Pr\left(\Delta b = inc \mid t_1 > t_2\right) = \Pr\left(\Delta b = inc \mid t_1 < t_2\right)$	✔			✔
$\text{T4}: cdf\left[\Delta b \mid t^R \text{low}\right] = cdf\left[\Delta b \mid t^R \text{high}\right]$	✔			
$\text{T4'}: \Pr\left(\Delta b = inc \mid t^R \text{low}\right) = \Pr\left(\Delta b = inc \mid t^R \text{high}\right)$	✔			✔
$\text{T5}: \Pr\left[\varepsilon_{1,j} > \varepsilon_{1,k} \mid \varepsilon_{2,j} > \varepsilon_{2,k} \,\&\, \min\left(\varepsilon_{1,j}, \varepsilon_{1,k}\right) > \max\left(\varepsilon_{2,j}, \varepsilon_{2,k}\right)\right] = \frac{1}{2}$			✔	

3. Application to eBay data

3.1 Data

We have three datasets at our disposal that include observations of $\left(t_1, t_2, b_1, b_2, inc\right)$. The datasets capture bidding on popular MP3 players in 2001, popular DVD's in 2002, and cars in 2003. Each dataset was selected to not end by the buy-it-now (BIN) option, and to have at least two bidders who bid above the minimum bid established by the seller (and above the reserve price, wherever applicable).[6] Ties are resolved as follows: b_1 is defined as the winning bid, so a tie at the top $(b_1 = b_2)$ implies $t_1 < t_2$. A tie for second highest bid is resolved in favor of the bid that would have won in the absence of the highest bid, i.e. t_2 is the time of the earliest second highest bid.

The MP3-player dataset captures 6122 auctions. For each auction, we observe the seller reputation, auction-characteristics like "photo included", the brand and model of the product sold, and whether or not the player was advertised as "new" by the seller. One player was particularly popular on eBay during the time of the data – the Diamond Rio 500 with 1370 listings. The second most popular player (KB Jamp3) only had 518 listings. The movie dataset captures 3555 auctions. For each auction, it includes an indicator of whether or not the seller was a "top seller", the title of the movie, and whether or not the DVD was advertised as "new" by the seller. The most popular movie in our sample is Black Hawk Down, with 848 listings.

The car data captures thousands of used cars sold in 2003 as well as a longer series on one of the most popular cars – the C5 Corvette. The Corvette data captures over 1300 auctions between 2001 and 2003, with exact timing available only for the 2003 data. The C5 is the fifth version of the Corvette - a version that is relatively new and relatively homogenous (produced without major changes from 1997 to 2004). To get the observable attributes of each car, we selected the auctions with a valid Vehicle Identification Number (VIN). A valid VIN gives information about the car including make, model, year, engine type, and model style, corroborating the car-information provided by the seller. We eliminated several observations, in which the VIN information did not agree with the information provided by the seller. Finally, we eliminated about two dozen auctions that sold for less than $10,000 or more than $40000 because these were outliers on the log scale (the median price was $26,200, std. deviation about $8,000).

These three datasets span a wide range of products sold on eBay, and they all contain the information on $(t_1, t_2, b_1, b_2, inc)$ because they were provided directly by eBay (in data obtained from the eBay website, b_1 would not be available). This information allows us to apply the tests that distinguish between the ascending and sealed models of bidding. We conduct these tests in two stages: First, we use tests T1-T4 that do not rely on the strong A3 assumption, and we can convincingly classify about 30 percent of the auctions as "not sealed" because of Δb being "too

[6] This selection obviously biases any demand measurements we make. It does not bias our testing of the sealed-bid abstraction as long as the bidders in the selected auctions do not bid systematically differently.

small", i.e. one increment or less. Using test T5 on all the data would thus have a foregone negative conclusion, but it is still possible that the remaining 70 percent of plausibly sealed "OverInc" auctions could be used to identify demand. We therefore apply test T5 only to the OverInc subset of auctions in the second stage.

Table 2:

Results of tests T1-T4

	MP3 players			**DVDs**			**Cars**		
	All	Rio 500	KB Jamp3	All	Black Hawk Down	Beaut. Mind	All Chevy Fall 03	Ford F150 Fall 03	C5 Vette 03
All auctions									
Number of observations	*6122*	*1370*	*518*	*3555*	*848*	*367*	*4955*	*510*	*642*
T1: Pr($t_1 > t_2$)	**64.5**	**61.5**	**65.1**	**72.6**	**72.3**	**72.2**	**65.2**	**68.8**	**60.3**
T2: Pr($\Delta b = inc$)	**14.4**	**12.4**	**19.5**	**14.3**	**14.2**	**17.2**	**13.0**	**14.1**	**18.1**
T3: ($\Delta b \mid t_1 > t_2$) vs. ($\Delta b \mid t_1 < t_2$)	56.8	26.7	97.3	**3.43**	**3.75**	22.2	**0.00**	10.9	0.03
T3': Pr($\Delta b = inc \mid t_1 < t_2$)	**4.46**	**1.91**	**8.07**	**5.45**	**4.68**	**7.84**	**5.32**	**4.55**	**10.6**
T3': ΔPr($\Delta b = inc$ | ">" vs. "<")	**14.8**	**15.5**	**17.1**	**12.1**	**13.1**	**12.9**	**11.5**	**12.5**	**12.4**
T4: ($\Delta b \mid t^R$ high) vs. ($\Delta b \mid t^R$ low)	**0.00**	**0.00**	**4.55**	**2.73**	**3.52**	97.0	16.1	30.9	6.89
T4': Pr($\Delta b = inc \mid t^R$ high)	**13.0**	**9.9**	**18.1**	**13.1**	**14.9**	**15.8**	**13.1**	**14.1**	**18.4**
T4': ΔPr($\Delta b = inc$ | low vs. high)	**2.73**	**4.93**	**2.70**	**2.41**	**-1.42**	**2.63**	-0.19	0.00	**-0.62**
OverInc ($\Delta b > inc$)									
Number of observations	*4338*	*953*	*361*	*2406*	*571*	*222*	*3646*	*352*	*452*
T1: Pr($t_1 > t_2$)	**64.8**	**62.4**	**64.3**	**73.2**	**74.4**	**73.4**	**62.5**	**67.6**	**57.1**
T3: ($\Delta b \mid t_1 > t_2$) vs. ($\Delta b \mid t_1 < t_2$)	**0.00**	**0.46**	23.1	**0.74**	**2.91**	34.5	**0.00**	**4.78**	**0.11**
T4: ($\Delta b \mid t^R$ high) vs. ($\Delta b \mid t^R$ low)	**0.00**	**0.00**	**0.06**	53.8	25.6	90.3	17.7	68.7	26.0
HighFirst ($t_1 < t_2$ & $\Delta b > 0$)									
Number of observations	*1885*	*418*	*161*	*972*	*235*	*102*	*1429*	*132*	*255*
T4: ($\Delta b \mid t^R$ high) vs. ($\Delta b \mid t^R$ low)	**0.00**	**0.16**	30.2	17.4	84.6	31.8	31.8	72.8	11.2
T4': Pr($\Delta b = inc \mid t^R$ high)	**3.51**	**0.48**	**6.25**	**4.73**	**4.27**	**7.84**	**6.16**	3.03	**10.2**
T4': ΔPr($\Delta b = inc$ | low vs. high)	**1.88**	**2.83**	**3.63**	**1.44**	**0.81**	0.00	**-1.69**	3.03	0.70

Note to Table 2: **Bold** values are rejections of the sealed-bid abstraction significant at the 5 percent level. All numbers are probabilities or differences in probabilities scaled between 0 and 100. T1, T2, T3' and T4' entries show the test-statistics (themselves probabilities), while T3 and T4 entries show the *p*-values of the Wilcoxon-Mann-Whitney rank-sum test.

3.2 Results of tests T1-T4 based on continuity of *F* or on the independence of timing

Table 2 documents the results of tests T1-T4 by dataset, and within each dataset by two popular products. Three tests reject the sealed-bid abstraction consistently across all datasets, with surprising empirical regularity of the test-statistics. First (T1), b_1 comes after b_2 in about two thirds of the auctions. This result suggests some but not prevalent ascending bidding. Second, about fifteen percent of the auctions end with the two bids exactly one increment apart (T2). This again suggests ascending behavior. In Section 4, we will show that the empirical $\Pr(\Delta b = inc)$ is "too high" to come from sealed bidding even when we relax the A2 assumption to allow mass-points at whole dollars and arbitrary multiples of the increment. Further suggesting that some auctions are better captured by the ascending model, the exactly-increment-apart outcome is about three times more likely when b_1 comes after b_2 compared to the reverse order (T3'). Therefore, the ascending-like timing tends to co-occur with the ascending-like Δb. Note that these uniformly-rejecting tests are based on different assumptions, and at least one of them remains valid whenever A1 or A2 holds.

The remaining tests reject the sealed-bid abstraction in some but not all datasets. With only a few exceptions, test T4' suggests that auctions with late bidding are about 2 percent more likely to involve $\Delta b = inc$. With the exception of the car data, test T4 finds larger Δb in auctions with early rather than late bidding. The late-bid auctions include all auctions, in which both of the top two bidders sniped. Therefore, the T4 result suggests that even sniping bidders react to the early bids, and even the snipe bids do not conform to the sealed-bid abstraction. The only test that does not produce consistent rejections is T3 – there does not seem to be a systematic effect of relative timing on Δb. But since exact ties in the top two bids involve $t_1 < t_2$ by definition, the subsample of $\{\Delta b \mid t_1 < t_2\}$ involves a substantial mass-point at zero. When exact ties are eliminated from the T3 comparison, the test rejects sealed bidding consistently as discussed next.

Tests T1-T4 cast serious doubt on the applicability of the sealed-bid abstraction to any of the three datasets because of mass-points of Δb at zero and *inc.* The auctions corresponding to those mass-points are obviously better captured by the ascending model. The question remains whether the remaining approximately 70 percent of the auctions with $b_1 > b_2 + inc$ fare better on the tests. If yes, then nonparametric demand identification ala Song would be possible at least in this "OverInc" subset of the data. We therefore reapply the tests that do not rely on inc (T1, T3, and T4) to the OverInc subset of the data. The results are also shown in Table 2, and they are quite discouraging: First (T1), not only does $\Pr(t_1 > t_2)$ remain significantly above ½, restricting attention to OverInc subset does not seem to reduce it at all. Second, as suggested above, test T3 now rejects the sealed-bid abstraction in most of the datasets. The emergent empirical regularity based on a Hodges-Lehmann estimate is that Δb is smaller when the highest bid follows after the second-highest bid compared to the other order (not reported). Finally, test T4 rejects the sealed-bid abstraction in the MP3 player data. In summary, based on the joint distribution of timing and bids, the sealed-bid abstraction seems doubtful even within the OverInc subset.

There is another subset of the data in which ascending bidding should not operate, namely $t_1 < t_2$: if the higher proxy bid is placed before the lower proxy bid, the high bidder probably did not react to b_2 in submitting his bid b_1. We call this subset "HighFirst", and we exclude $b_1 = b_2$ observations from it as well even though these involve $t_1 < t_2$ by our definition. We can apply tests T2, T4 and T4' to test the HighFirst data. The results of T2 are reported on Table 2 as part of the T3' test on all the data, and the size of the mass-point $\Delta b = inc$ is significantly bigger than zero but smaller than in full data (around 5 percent compared to 15 percent in MP3 players & DVSs, about 10 vs. 20 percent in cars). Test T4 rejects less often, but test T4' still rejects in most datasets. In summary, the sealed-bid abstraction is still significantly rejected in the HighFirst subset of auctions, but it is rejected less strongly and less consistently than in the OverInc subset or in the full data.

While tests T1-T4 reject the sealed-bid abstraction based on the joint distribution of (t_1, t_2, b_1, b_2) even in OverInc and HighFirst subsets, it is still possible that a carefully-selected subset of the (b_1, b_2) observations could be used in demand-identification. From the test-application perspective, it is also interesting to see how one would test the sealed-bid assumption without the timing data. Of the two subsets described so far, we focus on the OverInc subset because it contains a lot more data than the HighFirst one. In the next section, we apply the more fine-grained test T5 to the OverInc data to investigate the applicability of the sealed-bid abstraction without the timing and continuity assumptions.

3.3 Results of the tail-comparison test T5 based on the conditional-iid assumption

The first battery of tests rejected the sealed-bid model as a general property of all auctions. Specifically, about 30 percent of the auctions are suspect because the top two bids were exactly one increment apart or less. However, there remain 70 percent of the auctions with top two bids more than an increment apart, and it is possible that the sealed-bid model fits these auctions well-enough to permit nonparametric demand estimation based on conditional order-statistics. Test T5 is ideally suited to investigate this possibility.

As explained in Section 2.3, the first step of test T5 is conditioning on auction-level observables. To give the assumption A3 the best chance of holding, we first focus on one popular product at a time, and then we run the control regressions of remaining auction-level observables within that product. In the MP3-player and DVD-movie data, we focus on the two most popular products (Diamond Rio 500 and KB Jamp3 in players, Black Hawk Down and Beautiful Mind in movies). To further reduce unobserved heterogeneity, we then also further restrict the data on the most popular products to the listings stating "new" in the description. In that car data, we focus on the C5 Corvette. Note that in the case of C5 Corvettes, we have more observations for test T5 (732) than reported in Table 2 (452) because our dataset contains bid-data but no timing data from two additional years.

Different observables are available in the three different datasets, Table A2 in the appendix shows both the variables and the results of the control-regressions. As explained in Section 2.3, there are two ways to run each regression, we report the results of T5 using both approaches as well as using the raw bid data without any conditioning. For C5 corvettes, the theoretically preferable truncated regression of $\log b_{1,j}$ on Z_j truncated at $\log b_{2,j}$ did not converge, so we only report the results based on the OLS price-regression and the raw data.[7] The control regressions, if actually valid as demand-estimates, suggest several interesting properties of eBay bidding. Note that the OLS regression lumps together the effects of the explanatory variables on both entry to the auction (which is unobserved in general) and on the bids (valuations) of the bidders. The truncated regression, on the other hand, is conditional on entry and only measures the effects on valuations. It is thus interesting to observe that seller reputation and number of observed bidders are both strongly positively correlated with price in the OLS regressions, but not in corresponding truncated regressions. This is consistent with a model, in

[7] The lack of convergence is an early sign of problems with the sealed-bid abstraction and/or the iid assumption.

which reputation does not have a direct effect on the valuation of the good itself, but it increases prices by increasing entry into auctions of high-reputation sellers. These results, if valid, shed light on the underlying process behind the commonly observed effect of seller reputation on prices (Resnick et al. 2006). Analogously, more observed bidders obviously lead to a higher price, but observing more bidders does not make each bidder bid more on the good. Therefore, it seems that the "herding effect" in which bidders seek auctions with other bidders already present (Dholakia & Soltysinski 2001) is not due to those bidders valuing the goods more because of the presence of others. Conversely, observing more bidders does not make each bidder bid less as would be predicted by a common-value model, so one might conclude that the categories we observe are better captured by a private-value model. Of course, the above stories about demand suggested by the control regressions are only strictly valid when the tail-comparison test T5 does not reject the residuals. Table 3 shows that the test actually does reject the validity of these regressions.

Table 3: Results of the tail-comparison test (T5) on the OverInc subset of auctions

		Truncated regression residuals		*OLS regression residuals*		*Raw bid data*	
Product	Number of auctions	All pairs	Pairs with similar price	All pairs	Pairs with similar price	All pairs	Pairs with similar price
Diamond Rio 500 new	798	**62.0%**	**62.9%**	**61.9%**	**62.9%**	**57.4%**	**58.7%**
Diamond Rio 500	938	**60.9%**	**62.3%**	**60.6%**	**62.0%**	**55.8%**	**57.7%**
KB Jamp3	360	**60.5%**	**61.3%**	**59.2%**	**60.3%**	**54.6%**	**56.2%**
Black Hawk Down new	188	**58.9%**	**58.9%**	**58.7%**	**58.8%**	**58.8%**	**58.7%**
Black Hawk Down	570	**63.1%**	**62.4%**	**62.7%**	**61.9%**	**61.4%**	**60.8%**
Beautiful Mind	222	**65.9%**	**65.8%**	**66.6%**	**66.2%**	**65.5%**	**65.3%**
C5 Vettes	732	*did not converge*		**40.8%**	**44.4%**	**40.7%**	**41.2%**

Note to Table 3: The OverInc data includes all observations such that $b_1 > b_2 + inc$. All test-statistics are significantly different from ½ at 5% level . "Pairs with similar price" are all the pairs of auctions such that : $\left|\hat{\varepsilon}_{2,j} - \hat{\varepsilon}_{2,k}\right| < std\left(\hat{\varepsilon}_2\right)$.

In the *OverInc* subset of the data, the T5 test-statistic is significantly different from ½. Interestingly, the correction for observables does not make much difference in the test statistic.

Therefore, T5 rejects the sealed-bid abstraction for every product in every category under study, whether or not we control for observables. The patterns differ qualitatively across categories: in the feasible pairs of MP3-player and DVD-movie auctions, higher prices are correlated with higher first bids (T5 test-statistic is about 60%). In contrast, feasible pairs of car-auctions exhibit a negative dependence with a T5 test-statistic of about 42%. Another difference is in our ability to find feasible auction pairs in the first place: of all the auction-pairs, about 30 percent are feasible in MP3-players and DVD-movies, but only about 10 percent are feasible in cars (detailed analysis of feasibility not reported).

The reason behind the pattern in the Corvette C5 data is clear: the joint distribution of (b_1, b_2) is so concentrated near the "diagonal" $(\text{near } b_1 = b_2)$ that it is very difficult to find feasible pairs of auctions, and the tails of the conditional distributions of $(b_1 \mid b_2)$ are so steep that within the feasible pairs, the higher price makes a lower top bid more likely. In the DVD movies and MP3 players, the tails of $(b_1 \mid b_2)$ are actually quite flat, so higher prices are associated with higher top bids. The implied positive correlation of signals opens up the alternative explanation of auction-level demand-shocks unobservable to the econometrician: it could be that higher prices are associated with higher top bids because bidders observe some vertically-differentiating attribute not included in the control regressions. One way to check for this explanation is to further restrict attention to feasible pairs with similar prices, or to only new products that are presumably more homogeneous. Table 3 shows that the test-statistic is not consistently closer to ½ for new-product subsets of the data or for the "similar-price" subsets of the data. This stability of the test-statistic across more homogeneous subsets suggests that the departure from ½ is not due to auction-level demand-shocks unobservable to the econometrician. We will explore this alternative explanation further in the Robustness section.

Another way to use the idea behind test T5 would be to split the data on b_2 and run the truncated control-regression separately for low b_2 and high b_2. In data that satisfies A3, there should be no difference in the parameter estimates. Unfortunately, splitting the already-small samples results in mostly insignificant parameters, so it is difficult to find a significant difference even if there is one in the data-generating process. One example we did find is a differential time-trend among the MP3 players: based on above-median b_2, neither of the top two players exhibits a significant decline in valuation from February to April. However, based on below-median b_2 , there is a consistent and significant 1.5 percent decline. Another example is available from the DVD data: based on above-median b_2, the baseline valuation (the intercept of the regression) of Black Hawk Down is significantly lower than that based on below-median b_2. We do not report on these split regressions here, please contact the corresponding author for details. Note that it is a tall order for a truncated regression to pin down even the intercept regardless of selective sub-sampling, but a sealed-bid dataset that satisfies A3 must be able to fulfill this order.

4. Robustness checks

The biggest weakness of the fine-grained tail-comparison test (T5) is the assumption (A3) of no auction-level demand-shocks unobservable to the econometrician. Suppose instead that the regression equation is $\log\beta\left(x_{i,j}\right)=\alpha Z_j+\xi_j+\varepsilon_{i,j}$, and the econometrician does not observe ξ_j. Then, the test will reject spuriously whenever the variance of ξ_j is similar to the variance of $\varepsilon_{i,j}$. Consider the case of Apples and Bananas when the econometrician only observes fruit being auctioned. Suppose the auction is actually sealed, there are no observable differences Z_j, apples are privately valued at $\$A+Normal\left(\$0,\$1\right)$ and bananas are valued at $\$B+Normal\left(\$0,\$1\right)$. When $A\approx B$, apples are just like bananas for the purposes of the test. At the other extreme, when $A>>B$, apples and bananas populate separate regions of the support of $\left(b_1,b_2\right)$ and the test performs fine. Finally, when $A\approx B+1$, apples and bananas overlap in the data and conditioning on a lower b_2 confounds conditioning on the fruit being a banana with conditioning on the second highest bidder having a relatively low draw from the valuation distribution. The question thus arises: are the rejections in Table 3 due to unobserved shocks ξ_j or due to the sealed-bid abstraction not holding?

Two pieces of evidence against this alternative explanation are already contained in Table 3: First, the results on C5 Vettes cannot be explained by unobserved heterogeneity of products because unobserved ξ_j imply a positive correlation between the top two bids, i.e. T5 statistic above ½ . Second, T5 is not consistently closer to ½ when we consider more homogeneous subsets of the products (i.e. unused products or only auction pairs with similar prices). In this section, we provide a third piece of evidence against the alternative explanation by further restricting our attention to the HighFirst$\left(t_1<t_2\right)$ subset of the OverInc data. We consider HighFirst auctions to be “more sealed” both theoretically and empirically given T1-T4 (Table 2).

While the HighFirst auctions are less suspect of ascending bidding, they should be equally affected by unobservables. Table 4 shows that this is not the case: in 19 out of 20 of the product-regression combinations, the test-statistic is closer to ½ in OverInc & HighFirst data (Table 4) than in OverInc data (Table 3). The average change in the statistic is 2% towards ½ from an average of 11% away from ½. Incidentally, not all test-statistics in Table 4 are significantly different from ½ - a further sign of weaker evidence against the sealed-bid abstraction. We conclude from all three pieces of evidence that unobservable heterogeneity alone cannot alone explain the overwhelming rejection of the sealed-bid abstraction in *OverInc* data. That is: the rejection does not seem to be due to unobservable characteristics in the different auctions, but rather due to the fact that the sealed-bid abstraction does not hold.

Table 4: Does the tail-comparison test reject because of unobservables? Results of the tail-comparison test (T5) on the *OverInc & HighFirst* subset of auctions

		Truncated regression residuals			*OLS regression residuals*			*Raw data*	
Product	Number of auctions	All pairs	Pairs with similar price		All pairs	Pairs with similar price		All pairs	Pairs with similar price
Diamond Rio 500 new	240	**57.9%**	**58.7%**		**60.4%**	**60.8%**		**56.2%**	**56.7%**
Diamond Rio 500	278	**58.1%**	**59.0%**		**59.5%**	**60.5%**		54.0%	**55.3%**
KB Jamp3	114	**58.2%**	**59.6%**		55.6%	**57.2%**		53.3%	55.3%
Black Hawk Down new	55	58.6%	57.1%		57.5%	56.4%		59.6%	60.1%
Black Hawk Down	146	**62.0%**	**61.5%**		**61.2%**	**60.6%**		**60.5%**	**60.9%**
Beautiful Mind	59	**60.3%**	**61.2%**		**64.1%**	**62.5%**		**62.5%**	**62.8%**
C5 Vettes	215	*did not converge*			41.8%	50.0%		44.8%	44.8%

Note to Table 4: The OverInc & HighFirst data includes all observations such that $b_1 > b_2 + inc$ and $t_1 < t_2$. The test-stats significantly (at 5% level) different from ½ are shown in **bold**. "Pairs with a similar price" are all the pairs of auctions such that $\left|\hat{\varepsilon}_{2,j} - \hat{\varepsilon}_{2,k}\right| < std\left(\hat{\varepsilon}_2\right)$.

Another assumption that may be too strong for the reality of eBay is the continuity assumption A2: if many bidders bid in whole-dollar or whole-increment amounts, the resulting mass-points in the distribution of Δb may arise even in sealed bidding. The movie-data is ideally suited to check this possibility because the support of the bidding distribution only involves 20 whole-dollar amounts ($5-$25) , and the increment is 50 cents on the entire support, so there are only 19 additional whole-increment mass-points ($5.50, $6.50, …,$24.50). The empirical distribution of the highest bid indeed does not look continuous: 42.6 percent of the bids are whole dollars and additional 14.3 percent are a whole dollar plus 50 cents (considering the most popular movie – Black Hawk Down). To approximate the probability of *K*-increment differences $\Pr\left(\Delta b = K \cdot inc\right)$ one would expect from sealed bidding with such a lumpy distribution of bids, we perform a simple simulation of the top two bidders in a sealed-bid auction. We assume that there is a second-price sealed-bid auction with two bidders drawn iid from the empirical cdf of b_1 (and then ordered). By simulating a million repetitions of such an auction, we find that the expected mass-points are much lower than those observed in the data. We then increase the number of simulated bidders all the way to ten, keeping the distribution the same, and again examine the distribution of Δ*b*. The results this simulation exercise are shown in Table 5.

Table 5: Can a lumpy bid-distribution explain top two bids exactly increment apart?

Δb / *increment*	**Data**	Simulated sealed-bid auction, by number of simulated bidders								
		2 bidders	3 bidders	4 bidders	5 bidders	6 bidders	7 bidders	8 bidders	9 bidders	10 bidders
exactly 0	**4.37**	**2.64**	**3.66**	**4.29**	**4.79**	**5.12**	**5.47**	**5.82**	**6.08**	**6.46**
0-1	14.17	7.79	10.66	12.12	13.13	13.90	14.63	15.09	15.72	15.97
exactly 1	**14.17**	**2.69**	**3.43**	**3.68**	**3.98**	**4.14**	**4.17**	**4.21**	**4.23**	**4.32**
1-2	15.48	7.49	9.39	10.49	11.04	11.45	11.83	12.28	12.48	12.78
exactly 2	**8.74**	**4.16**	**5.14**	**5.59**	**5.83**	**6.07**	**6.18**	**6.18**	**6.38**	**6.47**
more than 2	43.07	75.23	67.74	63.83	61.23	59.32	57.73	56.42	55.10	54.00

Note to Table 5: The table shows both the actual and the predicted distributions of the difference between the top two bids Δ*b* measured in multiples of the *increment.* The distribution of bids used in simulations is the empirical distribution of b_1 for the Black Hawk Down movie, the increment is 50 cents throughout. The distribution of $\Delta b / inc$ is very similar for other movies as well as in the car and MP3-player data.

It is evident from Table 5 that mass-points in the distribution of signals may be able to explain the empirical probabilities of exact ties as well as the probability of the top two bids being exactly two increments apart. However, the expected $\Pr\left(\Delta b = inc\right)$ remains below 5 percent even for high numbers of simulated bidders – far below the empirical value of 14.2 percent. Therefore, a distribution of signals which is as lumpy as the distribution of the highest bid cannot explain the empirical probability of the top two bids being exactly increment apart. Before concluding, we need to discuss the implications of our findings for demand estimation – tone of the most managerially important applications of online auction data.

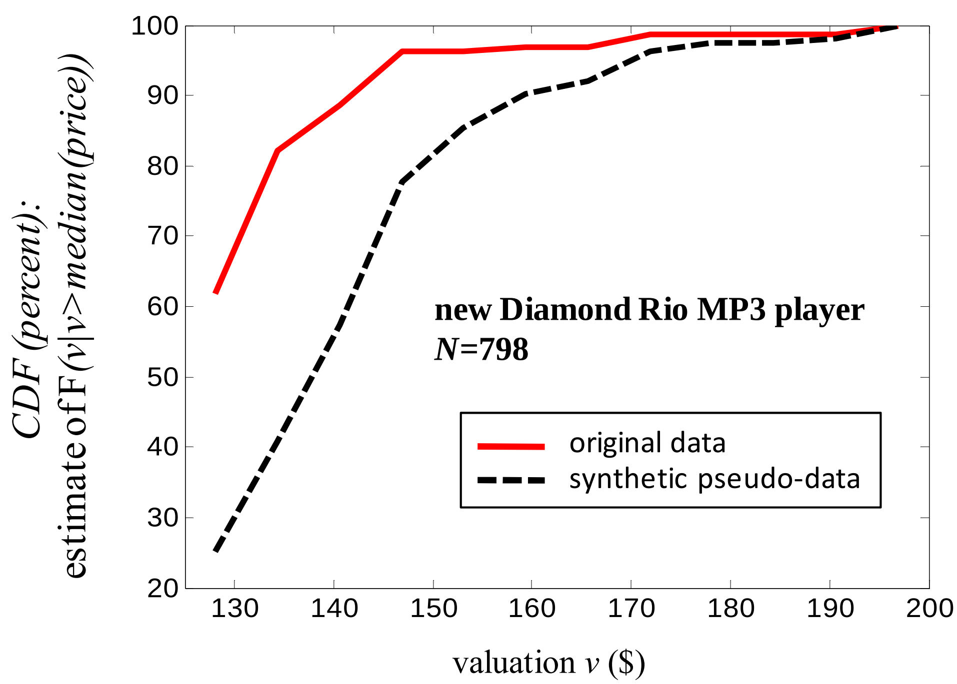
5. How can non-sealed behavior affect demand-estimates based on the sealed-bid abstraction?

Taken jointly, the tests T1-T5 show that eBay bidding does not satisfy the sealed-bid abstraction. This rejection implies a key question about the kind of biases one can expect from demand-estimates based on the abstraction. Specifically, what kind of bias can one expect when following Song (2004) and assuming that the $(\varepsilon_1, \varepsilon_2)$ data are generated with the likelihood of $\frac{f(\varepsilon_1)}{1-F(\varepsilon_2)}$? One way to answer this question would be to propose a new model of non-sealed bidding behavior, estimate demand using the new model, and compare the estimate to the estimate produced by the sealed-bid assumption. We cannot offer such a comparison in this paper because we are not proposing any specific new model. Instead, we assume the sealed-bid abstraction holds, and explore the potential bias by comparing the estimate arising from the actual data with an estimate on "sealed-like" pseudo-data based on the marginal empirical distributions of the top two bids.

The sealed-bid abstraction is rejected by the tail-comparison test (T5) because the high bids are essentially "too close" to their respective second-highest bids. This makes it difficult to find "feasible" pairs of auctions, i.e pairs (*k,n*) such that $\min(b_{1,k}, b_{1,n}) > \max(b_{2,k}, b_{2,n})$. For example, in the bidding data on new Diamond Rio 500 MP3 players, the percentage of auction-pairs that are feasible is 23 percent (raw data) and 29 percent (residuals of the truncated regression), while our simulations of the T5 test-statistic under ten different distributional assumptions consistently finds around 35 percent of pairs to be feasible in synthetic data that conforms to the sealed-bid abstraction.

Not only is it more difficult to find feasible pairs when b_1 is too close to b_2, the resulting demand-estimates have very thin tails. To illustrate one possible demand estimate, we ignore the auction-specific observables (in order to get a truly non-parametric estimate) and compute the implied empirical distribution of valuations above the median price for the new Diamond Rio 500 (N=798). This computation is quite simple – the estimate is a histogram of b_1 such that b_1>*median*(b_2) & b_2<*median*(b_2). Figure 2 shows that the estimated tail of valuations has most of the probability mass just above the median price of \$125: more than half of the valuations above \$125 are estimated to be below \$130. This estimate clearly lacks face validity.

Figure 2: Impact of non-sealed behavior on demand estimates



Note to Figure 2: Conditional cumulative distribution functions implied by two datasets: the original and a synthetic pseudo-dataset with the same number of observations and the same marginal distribution of b_1. The estimates are cumulative histograms with 12 equally-spaced bins of b_1 given b_1>*median*(b_2) & b_2<median(b_2). The percentage of observations such that b_1>*median*(b_2) & b_2<median(b_2) is 20% in the original data and 44% in the synthetic pseudo-data.

Figure 2 contrasts this unrealistic estimate with one guess of what a realistic estimate might look like: the dashed line is an estimate based on a synthetic dataset that matches the marginal distribution of b_1 but does not suffer from the "b_1 is too close to b_2" problem. We created each observation in the synthetic dataset by first drawing with replacement from the observed b_1, and then drawing with replacement from the observed b_2 until we found a b_2 that was actually lower than the respective b_1. We do not claim that the synthetic dataset magically de-biases the demand estimates. It could even bias them the other way. However, we suggest that the "estimate" based on our synthetic pseudo-data would be much more reasonable, and we argue that demand estimates based on the erroneous sealed-bid assumption will tend to be biased downward – towards distributions with very thin tails. This bias is bad news for pricing managers: the original hope for analysis of auction demand was to learn about the true valuations of the buyers, and perhaps alter pricing strategies accordingly. Our results suggest that such learning is not possible, at least not using the simple methods suggested by the sealed-bid abstraction.

What demand-estimation ideas survive our rejection of the sealed-bid abstraction? In terms of the structural parameters of the canonical auction model, we conclude that only the second order-statistic of the bidding distribution is observed on eBay (it is within an increment of the price); the highest proxy bid cannot be interpreted as the first order-statistic of the latent bid distribution. Instead, the highest bid is usually too close to the second-highest bid, and so when the residual of the second highest bid is unusually low, the residual of the highest bid is also unusually low and has a right tail that is too thin. Sadly, this makes the Song (2004) idea of conditional order-statistics less useful that we hoped it would be. Given that the number of bidders is unobserved even by eBay, standard demand-identification by the closing price and the number of bidders as discussed in (Athey & Haile 2002, 2005) is not possible either.

The only way to estimate demand while using only the second order-statistic is to somehow proxy for the number of bidders in each auction. We are aware of two approaches: Adams (2007) and Chan et al (2007). Adams (2007) shows that demand can be identified when the analyst observes a characteristic of the auction that is correlated with the expected number of bidders but uncorrelated with their item valuations – for example auction duration. Adams uses the Song (2004) structural model, which this paper suggests may be problematic. However, it is relatively straightforward to see that the structural assumptions in Haile and Tamer (2003) could be substituted for the structural assumptions of Song (2004) with all the identification results carrying through: Haile and Tamer (2003) assume that the auction is an open call ascending bid or English auction. The structural assumptions are that each bidder stays in the auction until the price rises above her maximum willingness to pay and each bidder keeps bidding as long as the price is below her willingness to pay. In a simple IPV "button" model with small enough increments, the price at the end of the auction is equal to the valuation of the second highest bidder. This is enough for the strategy proposed in Adams (2007) to estimate demand. Chan et al (2007) propose a way to account for the unobserved latent bidders directly within the Haile and Tamer (2003) bounds approach. They impute the latent bidders to be the bidders observed bidding on similar items at the time of the focal auction being analyzed. Given this imputation, it is possible to proceed as if the number of bidders were known, and the Haile and Tamer (2003) approach can be used to identify demand.

6. Discussion

We propose five different tests of the sealed-bid abstraction made by previous theoretical and empirical papers about online auctions. The tests rely on minimal assumptions, with different assumptions implying different tests. When applied to actual eBay data, the tests indicate that the sealed-bid abstraction does not describe eBay behavior well. This conclusion does not require all of the assumptions to hold: If it is reasonable to assume that timing of proxy bids is independent of their magnitudes, then we can reject sealed bidding because the highest bid tends to come after the second highest bid too often (about two thirds of the times), the difference between the top two bids is smaller when the highest bid comes after the second-highest bid, and the difference is between the top two bids is also smaller and when both bids arrive later in the auction. If it is reasonable to assume that the underlying distribution of valuations is continuous, then we can reject sealed bidding because about five percent of the auctions actually end in an exact tie (of the secret proxy bids), and about fifteen percent of the auctions end with the top two bids exactly one increment apart. These rejections cannot be explained by the actual distribution of valuations having lumps at whole-dollar and whole-increment amounts: based on our robustness check, simulated sealed bidding from the empirical distribution of bids would imply a dramatically lower chance of increment-apart observations than observed in the data. Finally, if it is reasonable to assume that there are no differences between the auctions observable to the bidders but not observable to the econometrician, then we can reject sealed bidding because the residuals of the potentially demand-identifying regression do not look like top two order-statistics of independent and identically distributed draws. Based on three robustness checks, this rejection cannot be explained by existence of unobserved differences between auctions. All of these negative findings replicate with surprising regularity across three different datasets spanning three years and diverse product categories.

Not only does the sealed-bid abstraction not fit the data in general, it does not fit even carefully selected subsets of the auctions in which one might expect it to hold. Specifically, we applied our battery of tests to plausibly sealed "OverInc" auctions that ended with the top two bids more than one increment apart, and to "HighFirst" auctions in which the higher bid was actually submitted first. Most of the tests rejected the abstraction in OverInc auctions, with the test-statistics barely changing compared to the full set of auctions in the data. The HighFirst auctions fared a little better, with only some tests rejecting in some datasets. Most notably, the chance that the top two bids are exactly one increment apart is about three times smaller among the HighFirst auctions.

These findings have profound implications for both empirical and theoretical analysts of online auctions. The empirical analysts cannot rely on the sealed-bid abstraction to identify demand ala Song (2004). We have shown that the data is sufficiently non-sealed to potentially produce large downward biases when the abstraction is used despite our evidence against it. Instead of providing the joint distribution of the top two order-statistics, eBay data seems to only contain the distribution of the second order-statistic (i.e. the price). The first order-statistic cannot be interpreted as the first order-statistic of the latent bid distribution. Without a new model that explicitly captures the timing of bids, the only information one can reliably glean from the data is that the winner's valuation was above the price, and that there was an unknown number of losers with valuations below the price (as in the bounds approach of Haile and Tamer 2003). We hope that our evidence spurs the development of new theoretical models that will better capture behavior and consequently allow for more fine-grained inference about the underlying structural primitives. In the rest of this Discussion, we summarize how our evidence constrains these theories.

While the sealed-bid abstraction fails, the data does not conform to purely ascending bidding either: the high bid does not always come after the second, and it is not always exactly one increment above. Therefore, our findings leave auction theorists with a puzzle: why would a bidder enter an online auction early and bid in some sort of ascending manner instead of waiting for the end and sniping? We can only speculate about the correct model here, and several possibilities emerge.

First, it is possible that bidders do not like to relinquish control over bidding to the proxy-bidding agent provided by eBay. This could be for purely psychological reasons, or because the bidder would like to keep the option of backing out of the auction at any moment. At the same time, however, coming back later in the auction is costly, so the bidder might as well put in some bid upon arriving at the auction. Therefore, the bidder may have to weigh the disutility of exposure (to paying a potentially high price) against the transaction costs necessary to remain completely in control (by committing to paying only an increment above the current highest price and likely having to return often). The size of this transaction cost of returning to the site determines which model fits eBay bidding-behavior best. When the transaction cost is so high that bidders will not return back to the same auction, Song's (2004) model arises, in which bidders bid on arrival and bidding is characterized by the sealed-bid abstraction. When the transaction cost is so small that bidders can return to the auction often, the ascending-bidding model arises and the sealed-bid abstraction is violated completely, at least in the private-value setting. In a common-value setting, the situation is a bit more complicated: Bajari & Hortacsu (2003) assume that the transaction cost of coming back is zero, and yet they show that it is never optimal to bid before the end of the auction. In their model, bidders do not care about being in control, and making the transaction-cost of coming back zero implies an equilibrium in which all bidders come back at the end, and bid in an effectively sealed-bid auction. Therefore, assuming something like bidders wanting to be in control is critical for the ascending-bid model to apply in the common-value setting. Finally, when the transaction-cost of coming back to the auction site is neither irrelevantly small nor prohibitively high, the bidders utilize the proxy-bidding agent at least somewhat to reduce the number of times they have to come back to the auction site. For example, the bidders may raise their proxy-bid by more than one increment, but not all the way to their secret maximum to avoid complete exposure.

A second theoretical possibility is that bidders are uncertain about their willingness to pay for the given item, but they can collect costly information to reduce their uncertainty. Rasmusen (2006) examined this possibility and concluded that since bidders want to incur the cost only if they are going to win, it is reasonable for a bidder to "increase his bid ceiling in the course of an auction" and react to other bidders. Whenever bidders face a transaction cost of returning to the auction, it is easy for them to submit a low-ball bid upon arrival, say a $10 on an MP3 player than usually sells for $200, and only incur the cost of finding out their true valuation if it looks like that particular auction is likely to conclude well below their valuation (about which they have an uncertain belief represented by a dispersed probability distribution).

Third, multiple bids by the same bidder may arise as the bidders consider multiple simultaneous auctions. Peters and Severinov (2006) obtain non-sealed bidding in an equilibrium of a model with sequential arrivals of bidders to many multiple simultaneous auctions. Nekipelov (2007) also focuses on multiple concurrent auctions, and shows that there is an incentive to bid early in order to deter entry by rival bidders. With incentives for early bidding balanced against the incentive to hide private information, within-auction dynamics that do not satisfy the sealed-bid abstraction emerge in Nekipelov's model. Finally, the incentive to hide private information may lead to ascending bidding even when the multiple auctions are sequential (as they are on eBay, ordered by their ending-times) instead of simultaneous: Cai et al (2007) argue that if eBay auctions were actually second-price sealed-bid auctions and eBay revealed all bids after every period, the sequential-auction game would have no symmetric pure-strategy equilibria: competitors would exploit revealed valuations in future rounds, so there is an incentive to hide one's valuation in the early rounds. Said (2008) argues that ascending auctions can reveal less information than sealed auctions, and provides a model of forward-looking bidding in sequential ascending auctions. Therefore, eBay bidders may bid in an ascending fashion to both hide their private information from competitors and to simplify their sequential-auction strategies. There are surely other explanations of the behavior we find, and we hope that future work resolves which of them is more appropriate as a model of eBay bidding.

Our results are applicable beyond eBay, in any market where both sealed and ascending bidding is made possible by the institutional details. For example, a procurement auction that runs live on the Internet but which also accepts proxy bids could be analyzed using our techniques. Another interesting area of application would be detection of collusion in sealed-bid auctions: if the top two bids in a sealed-bid auction are too related to each other, the auctioneer may conclude that the top two bidders are colluding.

References

Adams, Christopher (2007). Estimating Demand from eBay Prices. *International Journal of Industrial Organization* December 2007.

Adams, Christopher, Laura Hosken and Peter Newberry (2006). 'Vettes and Lemons on eBay. Federal Trade Commission. working paper, SSRN id 880780.

Ariely, Dan, Axel Ockenfels and Alvin Roth (2005). An Experimental Analysis of Ending Rules in Internet Auctions. *The RAND Journal of Economics* 36 (4), 890-907.

Athey, Susan and Philip Haile (2002). Identification in Standard Auction Models. *Econometrica* 70 (6), 2107-2140.

Athey, Susan and Philip Haile (2005). Nonparametric Approaches to Auctions. *Handbook of Econometrics, Volume* 6.

Bapna, Ravi, Wolfgang Jank and Galit Shmueli (2008). Consumer Surplus in Online Auctions. *Information Systems Research*, forthcoming.

Bajari, Patrick and Ali Hortacsu (2003). Winner's Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions. *RAND Journal of Economics* 34 (2), 329-55.

Bajari, Patrick and Ali Hortacsu (2004). Economic Insights from Internet Auctions. *Journal of Economic Literature* 12, 457-486.

Bradlow, Eric T., Young-Hoon Park (2007). Bayesian estimation of bid sequences in internet auctions using a generalized recordbreaking model. *Marketing Science* 26(2), 218–229.

Chakravarti, Dipankar., E. Greenleaf, A. Sinha, A. Cheema, J. C. Cox, D. Friedman, T. H. Ho, R. M. Isaac, A. A. Mitchell, A. Rapoport, M. H. Rothkopf, J. Srivastava, R. Zwick (2002). Auctions: Research opportunities in marketing. *Marketing Letters* 13(3) 281–296.

Cai, Gangshu., Peter R. Wurman and Xiuli Chao (2007) The Non-Existence of Equilibrium in Sequential Auctions when Bids are Revealed. *Journal of Electronic Commerce Research* 8, 141–156.

Chan, Tat Y., Vrinda Kadiyali, Young-Hoon Park (2007). Willingness to pay and competition in online auctions*. Journal of Marketing Research* 44(2), 324–333.

Dholakia, Utpal and Kerry Soltysinski (2001). Coveted or Overlooked? The Psychology of Bidding for Comparable Listings in Digital Auctions. *Marketing Letters* 12(3), 225-237.

Efron, Bradley and R.J.Tibshirani (1994) *An Introduction to the Bootstrap*. CRC Press, Boca Raton, FL.

Engelbrecht-Wiggans, Richard (1994). Sequential Auctions of Stochastically Equivalent Objects. *Economic Letters* 44, 87-90.

Haile, Philip A. and Elie T. Tamer (2003) Inference with an Incomplete Model of English Auctions. *Journal of Poltical Economy* 111(1), 1-51.

Milgrom, Paul R. and Robert J. Weber (1982). A Theory of Auctions and Competitive Bidding. *Econometrica* 50, 1089-1122.

Nekipelov, Denis (2007) Entry Deterrence and Learning Prevention on eBay, working paper.

Peters, Michael and Severinov, Sergei (2006) Internet auctions with many traders. *Journal of Economic Theory* 127(1), 220-245.

Rasmusen, Eric B. (2006). Strategic Implications of Uncertainty over One's Own Private Value in Auctions. *Advances in Theoretical Economics* 6(1).

Resnick, Paul, Richard Zeckhauser, John Swanson and Kate Lockwood (2006). The Value of Reputation on eBay: A Controlled Experiment. *Experimental Economics* 9, 79-101.

Roth, Al and Axel Ockenfels (2002). Last Minute Bidding and Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet. *American Economic Review* 92(4), 1093-1103.

Said, Maher (2008) Information Revelation and Random Entry in Sequential Ascending Auctions. Yale working paper.

Song, Unjy. (2004). Nonparametric Estimation of an eBay Auction Model with an Unknown Number of Bidders. working paper, UBC.

Yao, Song and Carl F. Mela (2008). Online Auction Demand. *Marketing Science*. forthcoming.

Yin, Pai-Ling (2006). Information Dispersion and Auction Prices. Harvard Business School Working Paper, No. 06-011.

Zeithammer, Robert (2006). Forward-looking bidding in online auctions. *Journal of Marketing Research* 43(3), pp. 462-476.

Appendix

Example of tail-comparison test (T5)

For illustration purposes, this example considers the raw Black Hawk Down data, focusing on the OverInc auctions. For all auctions *j,* the $\left(b_{1,j}, b_{2,j}\right)$ data been rescaled to fit between 0 and 1 by the transformation $b_{i,j} \rightarrow \dfrac{b_{i,j} - \min\limits_{j}\left(b_{2,j}\right)}{\max\limits_{j}\left(b_{1,j}\right) - \min\limits_{j}\left(b_{2,j}\right)}$. The table below shows the joint distribution of the rescaled $\left(b_1, b_2\right)$ with ten equal-sized bins. For example, the number 43 in b_2=0.25 and b_1=0.35 bin means that there are 43 $\left(b_1, b_2\right)$ observations such that b_2 is in the third decile and the corresponding b_1 is in the fourth decile: $b_2 \in (0.2, 0.3) \,\&\, b_1 \in (0.3, 0.4)$.

The outlined box contains a particular tail-comparison test that can be made. There are 153 observations in the box, and test T5 says that the distribution of these observations in any row within the outlined box should be the same. To increase the power of the test, let us compare the shaded row (96 observations) to all the rows above it. The (normalized) histograms of these two tails are shown below the table, and they are clearly not the same: the shaded row seems to have higher central tendency.

Table A1: Example of a joint distribution of $\left(b_1, b_2\right)$ with decile-sized bins

		b_1									
	0	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95
b_2	0.05	*2*	13	9	10	2	0	0	1	0	0
	0.15		*1*	20	12	5	3	0	0	0	0
	0.25			*9*	43	39	5	2	0	0	0
	0.35				*14*	**68**	**18**	**8**	**2**	**0**	**0**
	0.45					*30*	58	35	20	2	1
	0.55						*14*	50	11	1	0
	0.65							*18*	31	1	4

Histograms: Estimates of pdfs of tails							
	$b_1 \mid b_2 \in (0.3, 0.4) \,\&\, b_1 \geq 0.4$	71%	19%	8%	2%	0%	0%
	$b_1 \mid b_2 \leq 0.3 \,\&\, b_1 \geq 0.4$	81%	14%	4%	2%	0%	0%

Table A2: Control-regression results for all three datasets

		OLS regression of price		Truncated regression	
Dataset	Variable	Parameter	t-value	Parameter	t-value
Top player: new Diamond Rio 500 (N=798)	february	4.85	159.37	4.66	235.37
	march	4.7	165.7	4.55	28.25
	april	4.68	156.81	4.55	31.63
	log_sellrep	0.01	4.08	0.01	0.03
	photo	-0.03	-2.72	-0.04	-2.5
	bold	0.04	1.63	-0.04	-0.81
	gallery	0.00	0.06	-0.21	-1.44
	log(N_bidders)	0.02	2.4	-0.08	-0.66
		R^2	0.2	sigma	0.21
Top 2 models of MP3 players (N=1298)	rio	0.96	93.98	0.71	55.16
	february	3.87	160.03	3.86	81.57
	march	3.76	151.67	3.8	39.16
	april	3.73	147.54	3.83	39.36
	log_sellrep	0.01	3.44	-0.01	-0.1
	photo	-0.02	-1.99	-0.04	-0.42
	bold	0.02	0.91	-0.09	-2.24
	gallery	-0.00	-0.25	-0.06	-0.51
	log(N_bidders)	0.03	3.16	-0.04	-3.61
		R^2	0.89	sigma	0.21
Top movie: Black Hawk Down (N=570)	constant	1.94	58.55	1.81	96.64
	topseller	0.03	1.53	-0.04	-0.37
	new	0.15	6.97	0.06	1.04
	log(N_bidders)	0.11	5.1	0.09	1.53
		R^2	0.12	sigma	0.31
Top 3 movie titles (N=981)	constant	1.97	89.19	1.88	123.61
	Beautiful Mind	0.16	9.51	-0.04	-0.51
	Vanilla Sky	0.18	9.77	0.03	0.59
	topseller	0.05	3.79	-0.02	-0.38
	new	0.11	8.1	0.05	1.28
	log_Nbidders	0.09	6.39	0.06	1.31
		R^2	0.27	sigma	0.3

C5 Corvettes (N=732)	OLS regression of price	
Variable	Parameter	t-value
log_mileage	-0.012	-0.71
log_age	-0.272	-7.28
no_age	-1.019	-6.62
automatic	-0.026	-0.8
convertible	0.164	5
new	-0.749	-6.39
black	-0.072	-1.78
burgundy	-0.255	-3.17
red	-0.038	-0.92
silver	-0.088	-1.84
log_feedback	-0.022	-2.47
neg	-0.053	-0.77
year_2002	-0.211	-2.01
year_2003	-0.263	-2.52
month_feb	0.103	1.09
month_mar	0.192	2.29
month_apr	0.113	1.37
month_may	0.107	1.33
month_jun	0.123	1.5
month_jul	0.157	1.94
month_aug	0.228	2.78
month_sep	0.122	1.55
month_oct	0.12	1.53
month_nov	0.027	0.35
month_dec	-0.074	-0.92
num_bidders	0.009	3.32
constant	10.725	52.86
	R^2	0.52

Note to Table A2: Sigma is the standard deviation of the error term in the regression.The truncated regression did not converge in the C5 dataset .

Bootstrapping approach: Given the lack of analytical results about the distribution of the test-statistic under the null hypothesis, we used bootstrapping to test whether the observed magnitudes of T5 are significantly different from ½. To bootstrap the statistic given a sample of *N* auctions, we created 10,000 pseudo-datasets of *N* auctions by drawing from the original set of auctions with replacement. For each of the 10,000 datasets, we computed the statistic T5. We then reject the null whenever the confidence interval implied by the 2.5% and 97.5% quantiles of the 10,000 test-statistics does not include ½.

Pairs-of-pairs: properties of the proposed nonparametric test

To implement the tail-comparison test, we developed a nonparametric test that takes as input a joint distribution of top two order-statistics. This Appendix explores the properties of this test.

Test-development: Suppose there are *N* pairs of iid draws from some continuous distribution *F*. Denote the higher number in the *n*-th pair $y_{1,n}$ and the lower number $y_{2,n}$. Select all the pairs-of-pairs (*k,n*) such that $\min\left(y_{1,k}, y_{1,n}\right) > \max\left(y_{2,k}, y_{2,n}\right)$, and denote these the *feasible* pairs of pairs. From independence, it is clear that

$$\Pr\left[y_{1,k} > y_{1,n} \mid \min\left(y_{1,k}, y_{1,n}\right) > \max\left(y_{2,k}, y_{2,n}\right) \,\&\, y_{2,k} > y_{2,n}\right] = \frac{1}{2}$$

In other words, a relatively higher y_2 does not make a relatively higher y_1 more likely when both y_2's happen to be lower than both y_1's. This implies a test-statistic T5:

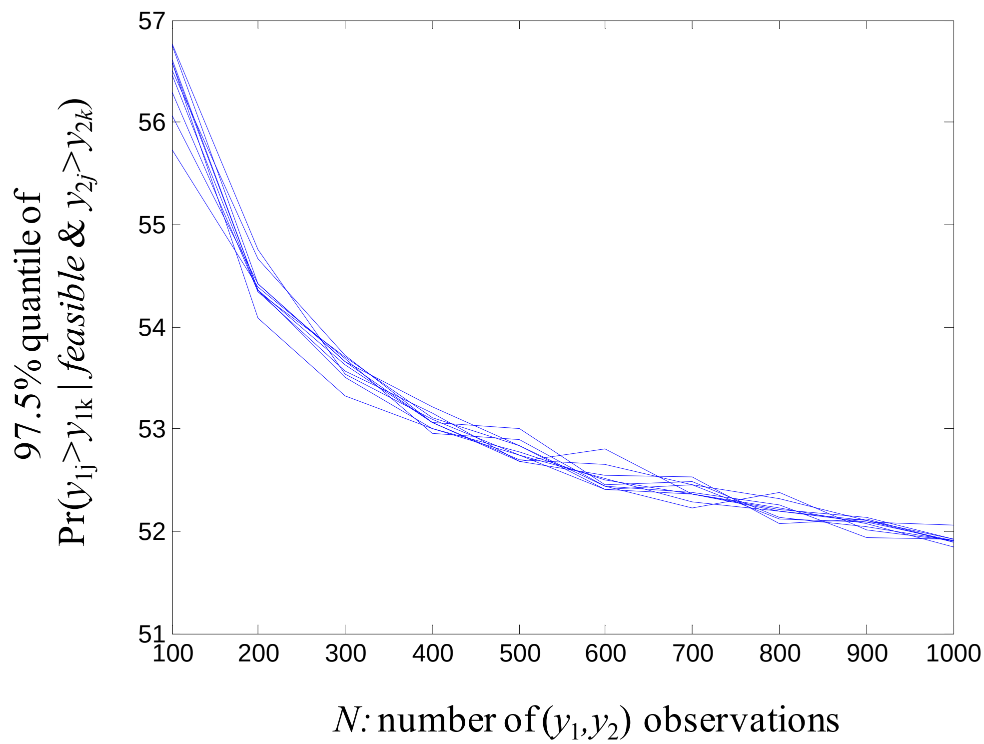
$$\mathrm{T5} = \frac{\sum_{j=1}^{N}\sum_{k=j}^{N} 1\left(\min\left(y_{1,j}, y_{1,k}\right) > \max\left(y_{2,j}, y_{2,j}\right)\right)\left[1\left(y_{1,j} > y_{1,k} \,\&\, y_{2,j} > y_{2,k}\right) + 1\left(y_{1,j} < y_{1,k} \,\&\, y_{2,j} < y_{2,k}\right)\right]}{\sum_{j=1}^{N}\sum_{k=j}^{N} 1\left(\min\left(y_{1,j}, y_{1,j}\right) > \max\left(y_{2,j}, y_{2,k}\right)\right)}$$

Test properties: The null prediction that T5=½ does not depend on *F,* but the distribution of the test-statistic around ½ is difficult to derive analytically because *F* does influence the chance of finding feasible pairs-of-pairs as well as the pattern of dependence across the pairs of pairs. To investigate the actual properties of the test-statistic, we simulated 1000 datasets with different *N* ranging from 100 to 1000, and with 10 different *F* (Normal, Lognormal, Gamma, Weibull, Beta,and Uniform, with different parameters). We found that the distribution of the test-statistic does not in fact seem to depend on *F* : it is approximately Normal even for *N*=100, and the 95% confidence interval shrinks down proportionally to $\frac{1}{\sqrt{N}}$. Please see Figure 3 below for a plot of the upper bound of the 95% confidence interval as a function of *N* : each line in the Figure represents one distribution, and it is clear that all the lines essentially coincide. These findings replicate even when $y_{1,n}$ and $y_{2,n}$ are actually the top two order-statistics of an arbitrary number of draws from *F*.

It is instructive to relate the confidence interval we found to a naïve theory one might have about the distribution of the test-statistic. If there happen to be *K* independent feasible pairs of pairs, then the total number of feasible pairs with $\left[y_{1,k} > y_{1,n} \mid y_{2,k} > y_{2,n}\right]$ would be distributed

$Binomial\left(p=\frac{1}{2},K\right)$, which would in turn be asymptotically approximated by $Normal\left(pK,p\left(1-p\right)K\right)=Normal\left(\frac{K}{2},\frac{K}{4}\right)$. Therefore, one might naively expect that the proportion of feasible pairs with $\left[y_{1,k}>y_{1,n}\mid y_{2,k}>y_{2,n}\right]$ would be approximately $Normal\left(\frac{1}{2},\frac{1}{4K}\right)$. We find that the variance of the order-statistic is in fact far greater than *K,* and it is related to *N* instead of *K.* Specifically, we regressed the size of the confidence intervals on $\frac{4*1.96}{\sqrt{4N}}$, and found the coefficient to be about $0.313\approx\sqrt{0.1}$ with $R^2=0.995$. This regression implies that the distribution of the test-statistic under the null hypothesis is extremely well approximated by $Normal\left(\frac{1}{2},\frac{0.1}{4N}\right)$. In other words, the amount of information in feasible pairs of a sample of *N* pairs can be approximated by the information that *would be* contained in 10*N* independent feasible pairs of pairs.

Figure 3: 97.5% quantile of the T5 test-statistic for 10 different distributions



Note to Figure 3: Each line corresponds to a different distribution. For each distribution and *N*, 1000 datasets were generated to compute the quantile.