Shill Bidding, Reserve Price and Seller's Revenue

Kong-Pin Chen* Ting-Peng Liang† Ted Chang‡ Yi-Chun Liu§ Ya-Ting Yu¶

August 15, 2018

Abstract

Starting with a reserve price has been shown in some empirical literature to be surprisingly effective in online auctions, in that it generates greater expected revenue than a reserve price which is consistent with theory. However, literature has also pointed out that sellers who practice shill bidding also start with a low reserve price. This raises the question of whether this empirical result is only an illusion mainly caused by shill bidding. Using data from eBay Motors, this paper empirically reinvestigates the effect of reserve price on the seller's sale rate, transaction price and expected revenue, taking into account the impact of shill bidding. We first construct a shill score for each listing, then estimate the trade probability and transaction price equations, while controlling for shill bidding. We find that shill bidding increases transaction price, a result consistent with its purpose. Moreover, increasing the reserve price increases transaction price but reduces trade probability. Despite this, reducing the reserve price still results in higher expected revenue. Finally, our empirical result rejects a recent theoretical prediction that shill bids mainly occur as last-minute bids.

*Research Center for Humanities and Social Sciences, Academia Sinica, Taiwan. E-mail: kongpin@gate.sinica.edu.tw.

†Department of Information Management, National Sun Yat-sen University, Taiwan. E-mail: liang@mis.nsysu.edu.tw.

‡Research Center for Humanities and Social Sciences, Academia Sinica, Taiwan. E-mail: f95323027@ntu.edu.tw.

§Taiwan Semiconductor Manufacturing Company, Taiwan. E-mail: gfm2234@gmail.com.

¶Research Center for Humanities and Social Sciences, Academia Sinica, Taiwan. E-mail: yyating@gmail.com.

1 Introduction

Although the function served by the reserve price in an auction has been solidly established in theory,[1] in practice its effectiveness is rather mixed. On the one hand, Ariely and Simonson (2003), Häubl and Popkowski Leszezye (2003), Reiley (2006), Brown and Morgan (2009), and Choi et al. (2016) show that an increase in reserve price usually decreases the number of bidders but increases the transaction price conditional on sale. On the other hand, Barrymore and Raviv (2009), Ku et al. (2006), Simonsohn and Ariely (2008), and Bajari and Hortaçsu (2003) find a negative effect of reserve price on revenue. Finally, Lucking-Reiley et al. (2007) and Einav et al. (2015) find no effect at all.

To reconcile theory with the empirical results, several explanations have been proposed. Milgrom and Weber (1982) and Levin and Smith (1996) show that when bidders' valuations are correlated, it might be optimal for a seller to set a low reserve price to encourage early bidding, thereby facilitating information revelation. McAfee and McMillan (1987) and Levin and Smith (1994) show that, when entry is costly for bidders, sellers might find it optimal to set a low, or even zero, reserve price to encourage entry. The third explanation, which is somewhat related to this paper, is to create an atmosphere of a bidding fever, so that the bidders are thrilled by the fervor of competing against each others, and on the way increase their willingness-to-pay beyond the level of what it normally would be. By setting a low reserve price for an item, the seller hopes to attract as many bidders as possible to promote bidding fever.[2]

One of the reasons for setting a low reserve price that has attracted little attention is sellers' shill bidding. Shill bidding is when a seller of an item has a third party bid, not to buy that item,

[1] For example, Myerson (1981) and Riley and Samuelson (1981) show that, for the independent-valuation case with no entry cost, the optimal reserve price is higher than the seller's valuation of the item.

[2] For relevant literature, see Heyman et al. (2004), Adam et al. (2015), Ku et al. (2005), Jones (2011), and Adam et al. (2011).

but solely with the intention to drive up its price.[3][4]

The literature has pointed out that sellers who shill often start with a low reserve price (e.g., Steiglitz, 2007; Kauffman and Wood, 2003). Under this practice, the seller deliberately sets a low starting price, attracting the bidders' attention to lure them into the auction. Once bids are placed on the item, the seller places phantom bids to compete with bidders, or even to create bidding fever. Given that the main reason to shill is to increase transaction price, an obvious but little investigated issue is whether the fact that a low reserve price often yields a higher revenue is only an illusion caused by shill bidding. Put differently, the relationship between the reserve price and the transaction outcome predicted by theory is garbled when the seller complements a low reserve price with shill bids. If we do not control for the possibility of shill bidding, a low reserve price will seem a very effective and rewarding practice for the sellers. Using data from eBay Motors, our first purpose is to empirically investigate how shill bidding affects the outcomes of the auctions. We then reinvestigate the relationship between the reserve price and auction outcomes, while controlling for the influence of shill bidding.

Shill bidding, however, is extremely difficult to detect. Although the history of shill bidding can be traced back to the 19th century (Engelbrecht-Wiggans and Nonnenmacher, 1999), it becomes much easier to practice in an online auction, and is much more difficult to detect. For one thing, we usually do not have the IPs, or even the full ID of the bidders.[5] Furthermore, the sellers who shill can create as many fake IDs as they see useful, making it difficult to identify

[3] For theoretical literature that shows how shill bidding increases a seller's revenue, see Graham et al. (1990), Delta(1999), Bag et al. (2000) and Izmalkov (2004) for the independent valuation model; and Vincent (1995) and Chakraborty and Kosmopoulou (2004) for the common value model.

[4] There are actually two reasons for shill bidding. The first, as defined above, is the seller's attempt to compete (either by a third party's or her own bids) with other bidders to drive up the transaction price. The second reason to shill, called reserve price shilling in the literature, is to avoid the fee paid by the seller to the auction platform. Since the platform usually charges higher listing fees for items with higher starting bids, in order to reduce the fees, some sellers set a low starting bid, and then place shill bids to raise bids to levels that they desire (Kauffman and Wood, 2003). However, for automobiles, eBay charges a fixed listing fee, plus another fixed listing fee for items that are sold (McGrath and McGrath, 2010). Therefore, fees are independent of starting bids, and there is no need to practice reserve price shilling. When we mention shill bid in this paper, we mean the first type.

[5] On the website, we usually only observe the first and last letters of a bidder's ID.

the bidders who consistently bid in the same seller's auctions but never win. Finally, even if we can observe the full ID, it is still difficult to detect shilling. This is because shill bidding has become something of an industry, where professionals place phantom bids on behalf of sellers.[6] Given that it is almost impossible to know who the shill bidders are for sure, it is very difficult to control for its influence.

Because of this difficulty, the literature has proposed several "shill-bid scores" to measure how likely it is that an online listing has been shilled, based on each bidder's bidding behavior, rather than IP. The scores are constructed by considering several regularities in a bidder's behavior who shills, and assigns weights to each of the regularities to construct a summary score for the likelihood that a listing has been shilled. Adopting this methodology, we constructed a shill-bid score for each listing. By controlling for the possibility of shill bidding using the shill score, we then reinvestigated the effect of reserve price on transaction outcomes. Before we can construct the shill score, however, we need to identify the bidders. This is because most of the bidders's IDs are usually partially concealed. By tracking the bidding behavior and the average daily change of experience scores of the bidders who fully reveal their IDs, we use an algorithm to identify the concealed bidders.

As expected, our results show that shilling increases the transaction price. Moreover, reserve price has a negative effect on trade probability, and a positive effect on transaction price conditional on sale. Finally, even after controlling for shill bids, there is still an inverse relationship between the reserve price and the seller's expected revenue. This implies that starting with a low price remains a good strategy for the seller, even absent shill bids.

There are three papers we are aware of that investigate the effect of shill bidding on auction outcomes. Kaffman and Wood (2005) is mainly concerned with reserve price shilling, i.e., shilling that aims to avoid listing fees. They show that reserve price shilling, in the form of premium

[6] See, for example, the discussion NamePros: https://www.namepros.com/threads/giant-shill-bidding-operation-at-namejet-exposed.1013479

bids (i.e., irregular bids that are higher than other bids for the same items in other auctions), increases the final prices. Therefore, even those shill bids which mainly aim to avoid fees can positively affect transaction prices. Kosmopoulou and DeSilva (2007) conduct experiments in a common-value setting to show that, if the bidders are aware of the possibility of shilling in an auction, then seller's revenue will actually decline because the bidders adjust their bidding strategies. Relatedly, Bose and Daripa (2017) theoretically show that, if the bidders suspect seller's shilling, then they will strategically snipe (i.e., place bids only in the last minute of the auction).[7] Therefore, the only chance for the sellers to shill is near the end of the auctions. This theory has the strong implication that shill bids occur mainly before the auction ends. With the shill score we construct, we can test this implication. Specifically, if the theory is correct, the shill scores of the bidders who place bids near the end of the auction should be substantially greater than those who place bids earlier. Using several measures for the "last minute" of the auction, we do not find any positive correlation between the shill score and the bidders who snipe. In fact, the correlation is significantly negative. Our result thus rejects the theory's prediction.

2 Data

The data were collected from the listings of Toyota cars in eBay Motors for a nine-month period from June 18, 2008 to March 6, 2009. There were 37,357 listings and 351,595 bid records. We first deleted 2,808 listings of new cars from the sample. There are two reasons for this. First, there is no Blue Book value for new cars. The Blue Book value is important for the empirical studies on the prices of used cars, as it is a good indicator for the value of used car based on their characteristics, and is widely consulted. In the price regression, it serves as a good proxy for many variables concerning a car's characteristics and conditions that affect its value. Second, not only do new cars account for less than 8% of our sample, but also only 147 of the 2,808 new

[7] For discussion of sniping, see Steiglitz (2007) and Ockenfels et al. (2006).

cars were sold (5% of new cars; while the sale rate for the whole sample is 19%). This implies that eBay Motors is predominantly a used-car platform. We further deleted observations that were posted-price listings and listings with best offer, as they were not really auctions.

Our empirical study consists of two related parts, using different samples. In the first part, since most of the bidders did not fully reveal their IDs, we had to identify the bidders. Since some sellers' IDs were missing during data collection, we deleted these listings from the sample, as there is no way to know whether two listings with concealed IDs are from the same seller. In all there remained 7,653 sellers in the sample (see Table 2). The bidder's ID, however, can be either fully revealed or partially concealed. Moreover, the seller can choose to conceal the IDs of the bidders in her listing. In that case, all the bidders' IDs in the listing will be completely concealed. For a partially concealed ID, we only observed its first and last alphabets (or symbols). Among the 181,819 bids that remained after we deleted new cars and non-auctions, 3,369 were from fully revealed IDs, 151,230 were partially revealed, and 27,220 were completely concealed.[8] Similarly, we deleted all bids from completely concealed IDs, together with any listing whose sellers concealed the bidders' IDs. That gave us 9,473 regular auctions and 8,968 buy-it-now auctions, and we used these 18,441 listings and 154,599 bids to identify the bidders, and then to construct the shill scores for the listings. The definition of variables and the summary statistics of the sample for the first part is reported in Tables 1 and 2.

In the second part of our empirical study, we investigate how shill bidding and the reserve price affect the transaction outcomes. On eBay, sellers are not allowed to set an open reserve price.[9] Therefore, the de facto reserve price for the seller is the starting bid, which the sellers must provide when they list an item (except for fixed-price listing). Hereafter, we will use the two terms "reserve price" and "starting bid" interchangeably. In this part, for the reason explained

[8] If we consider listings, rather than bids, there were 2,026 listings in which all bidder's IDs were completely concealed, 303 listings in which all bidders revealed their full IDs, 16,849 listings in which bidders' IDs were partially revealed, and 1,289 listings in which some bidders fully revealed and some partially revealed their IDs.

[9] The seller is only given the option to set a secret reserve price.

earlier, we used only the sample in which we could find the Blue Book values of the cars.[10] Also, the buy-it-now auctions in the sample deserve two explanations. First, the buy-it-now option on eBay is temporary, in the sense that if any bidder places a bid, rather than exercising buy-it-now, then the buy-it-now option disappears, and the buy-it-now auction reduces to a regular auction. Therefore, the items in the buy-it-now auctions can either be sold at the buy-it-now price (when a bidder exercises the option), or at the second highest bid (when a bidder places an eligible bid before buy-it-now is exercised). Since the former is essentially sold with a posted price, we deleted buy-it-now auctions which were sold at a buy-it-now price from the sample. Second, as shown in Chen at al. (2017), the seller's optimal reserve prices for different listing formats are usually different. In particular, the optimal reserve price for a buy-it-now auction is higher than that for a regular auction. Therefore, in the empirical model, we used a buy-it-now dummy to control for the difference in the reserve price between the two listing formats.

After deleting listings whose Blue Book values could not be identified, the buy-it-now listings that were sold at buy-it-now prices, and listings that had missing values, there remained 10,893 listings, of which 5,268 were regular auctions, and 5,625 buy-it-now auctions. For these listings, the bidder's IDs were completely revealed in 1,153 bids, and partially revealed in 68,766 bids. Therefore, there were 69,919 bids and, as we will see in the next section, 25,896 distinct bidders in the sample for this part.

For each listing, the data contain (i) the auction characteristics, including the starting price, the auction duration posted by the seller, whether there is a secret reserve price, whether the item is sold, and the transaction price if it is sold; (ii) car characteristics, including car age, mileage, vehicle model and body type, fuel type, etc.; (iii) seller's characteristics, such as whether the seller is a dealer or not, the seller's experience and feedback score. In addition, we also collected the bid history for each listing (for instance, bid amount and time of bid). Among the 10,893

[10] In our data, there are two main reasons that we could not find the Blue Book value of a used car. First, the seller did not give the car age. Second, we did not know whether the seller is a dealer.

listings, about 17.1% were sold, whose average ratio of transaction price to Blue Book price is 0.70. Table 3 and 4 report the variable definitions and their summary statistics for the sample of the second part. Table 5 summarizes the numbers of listings, bidders, and sellers in the two parts of the empirical study.

3 Empirical Model

In this section, we first construct the shill scores for the bidders and the listings, together with how we handle the problem of partially concealed IDs. Based on the listing's shill score, we then build up an empirical model to estimate the trade probability and the transaction price, taking into consideration the influence of shill bids.

3.1 Constructing the Shill Score

One convenient way to detect shill bidding is through the bidder's IP. The reasoning is that if the bid comes from the same IP as the seller's, then it must be placed by the seller.[11] This is not possible for our study because we could not track the bidders' or sellers' IPs. However, it is well known that shill bids are often placed by a third party (possibly professionals) on behalf of sellers. Even if we access the seller's ID, many shill bids remain undetected. Therefore, the literature also identifies shill biddings through an operational definition, i.e., by the bidder's bidding process and their outcomes, rather than their source. In that case, the first step for detecting shill bids is to identify the bidders and the sellers. Again, as explained earlier, the bidders' IDs are usually concealed.[12]

In order to identify the bidders, we adopt an identification method proposed in Liu (2017).

[11] For studies which trace user IP to detect shill bidding, see Mamun et al. (2013) and Mamun (2015).

[12] For example, a seller whose ID starts with b and ends with k is usually shown as b***k. No matter how long the ID is, there are always 3 asterisks in the middle.

First, a preliminary identification is made through checking the ratings of the bidders. Unlike the bidders' IDs, the ratings of the bidders are observable. There were 1,691 bidders in the sample who fully revealed their IDs, and 697 of them bid at least twice. Since we observe the bidders only when they bid, these 697 bidders were essentially the only full-ID bidders who we knew for sure appeared at least twice. We calculated the average daily change of ratings for these bidders, which was 0.12. Within the group of partially concealed IDs, if the change of the ratings of two IDs exceeded 0.12 per day, we viewed them as coming from two different bidders. Therefore, two bidders who had the same first and last alphanumeric characters in their (partially concealed) IDs were identified as the same bidder if and only if their ratings did not differ by 0.12 per day on average. We then classified bidders into groups, and bidders in the same group were regarded as the same bidder. The procedure produced 41,555 distinct bidders. To test how well this identification procedure works, we applied it to the subsample of bidders whose IDs were completely revealed, and found 1,508 distinct bidders (i.e., 1,508 groups), of which 1,313 were correctly identified by our criterion when we compared them to full IDs (precision rate: 1,313/1,691 = 78%).

Since the criterion that two IDs are classified as identical if the daily change of ratings is less than 0.12 is probably too soft,[13] we apply the Bayesian Information Criterion (BIC) to the grouped IDs to further identify them, based on the informational similarity of their bidding behavior. The information includes a bidder's reputation score, number of bids he placed, number items he won, time of bids, and bid increments. Specifically, for every bidder in every listing, we gathered information on his reputation score, the number of bids he placed, the length of time between his first and last bids, the lengths of time from his first and last bids until the end of the auction, and whether he won the item. Based on the information, we then compared the similarity of behavior between bidders who have two identical concealed IDs to further classify them into

[13] This can be seen from the fact that there were actually 1,691 bidders with full ID. The criterion only produced 1,508 distinct bidders.

smaller groups. With this procedure, we identified 50,994 distinct bidders. Therefore, together with the 1,691 bidders who fully revealed their IDs, we had 52,685 bidders in our sample. Again, we applied it to the subsample of bidders whose IDs were completely revealed, and identified 1,511 distinct bidders. Compared to full IDs, we correctly identified 1,369 bidders, with a precision rate of 1,369/1,691 = 81%. Though not perfect, we believe this is accurate enough to justify our identification procedure as a first step in constructing the shill score.

In total, there were 18,841 listings, 52,685 distinct bidders, and 7,653 sellers in the first part of the empirical study, and 10,893 listings, 25,896 bidders, and 4,433 sellers in the second part of the empirical study (see Table 5). When every bidder was assigned a distinct ID, we can investigate the likelihood that a bidder is a shill bidder through his bidding histories, and then construct the shill score of each listing. Kauffman and Wood (2003) identified shill bidders through their questionable bidding behavior. They reasoned that if a bidder chooses to bid in an auction when he has the chance to place the same or a lower bid in another concurrent auction featuring an identical item, and he does not bid in both auctions, then this bidder is likely to be a shill bidder. Shah et al. (2003) detected shill bidders through estimating how likely a bidder was to participate and win auctions held by different sellers. Xu, Bates and Shatz (2009) used multiple criteria on the behavior of bidders to check for shill bidding. The criteria included early bidding time, a large number of bids, and small bidding price increments. Dong, Shatz and Xu (2009) not only used various questionable bidder behaviors to identify potential shill bidders, but also improved and verified the detection model by applying the Dempster-Shafter theory.

In this paper, we adopt the method proposed by Trevathan and Read (2009), who constructed "shill score" for each bidder to measure how likely it is that he is a shill bidder. The underlying assumption of their detecting procedure is that a shill bidder should exhibit the following characteristics: (i) A shill usually bids exclusively in auctions of one particular seller; (ii) a shill tends to have a higher bid frequency;[14] (iii) a shill tends to have very few wins for the auctions participated in; (iv) a shill bid generally follows a new bid within a very short time; (v) a shill usually out-bids rivals by minimum increments; and (vi) a shill bid tends to occur at the beginning of the auction. Liu (2017) extends the approach of Trevathan and Read (2009).[15] Following their identification procedure, the shill score of bidder i in listing j in this paper is defined as:

$$Shill\ Score_{i,j} = \frac{\theta_1\alpha_{i,j} + \theta_2\beta_{i,j} + \theta_3\gamma_{i,j} + \theta_4\delta_{i,j} + \theta_5\varepsilon_{i,j} + \theta_6\zeta_{i,j} + \theta_7\eta_{i,j}}{\theta_1 + \theta_2 + \theta_3 + \theta_4 + \theta_5 + \theta_6 + \theta_7}, \tag{1}$$

where

$\alpha_{i,j}$ = the percentage of listings held by seller m participated in by bidder i.

$\beta_{i,j}$ = the relative number of bids bidder i placed in listing j.

$\gamma_{i,j}$ = the proportion of bidder i losing in listings held by seller m.

$\delta_{i,j}$ = the average time difference between bids placed by bidder i and the latest bid before that in listing j.

$\varepsilon_{i,j}$ = the average bidding price difference between bids placed by bidder i and the latest bid before that in listing j.

$\zeta_{i,j}$ = the time difference between the first bid placed by bidder i and the starting time in listing j.

$\eta_{i,j}$ = the time difference between the last bid placed by bidder i and the ending time in listing j.

In the equation, $\delta_{i,j}$, $\varepsilon_{i,j}$, $\zeta_{i,j}$, and $\eta_{i,j}$ are all relative differences compared to other bids placed by other bidders in each listing. Moreover, shill score is normalized in such a way that its value lies between 0 and 1. The greater the number, the more likely bidder i is to be a shill bidder in

[14] eBay shows two formats of bid history, one includes the automatic bids (i.e., the proxy bids submitted by eBay's automatic system on behalf of the bidders), and one without. Since automatic bids are not shill bids, but only a result of the bidder's submitting a relatively high bid, our data on bid frequency uses the count without automatic bids.

[15] For the details of constructing the shill scores, see Trevathan and Read (2009) and Liu (2017).

listing j.

The values of the parameters above were calculated from the sample, while for the values of the weights, the θ_i's ($i = 1$, ..., 7), Principal Component Analysis (PCA) was used to derive them. The PCA operation eventually yielded the following formula:[16]

$$Shill\ Score_{i,j} = \frac{2 \times \alpha_{i,j} + \beta_{i,j} + \gamma_{i,j} + 3 \times \delta_{i,j} + \varepsilon_{i,j} + 3 \times \zeta_{i,j} + 3 \times \eta_{i,j}}{14}. \tag{2}$$

After defining the shill score of each bidder in a listing, the average shill score of each listing j is its simple average:

$$Average\ Shill\ Score_j = \frac{1}{n_{i|i\in j}} \sum_{i\in j} Shill\ Score_{i,j}, \tag{3}$$

where $n_{i|i\in j}$ is the number of bidders participating in auction j. The $Average\ Shill\ Score_j$ is meant to portray how likely it is that listing j is being shilled, and was used in our model to control for the effect of shill bidding. Again, the value of the shill score for an auction lies between 0 and 1.

The distribution of the bidder's shill scores is plotted in Figure 1, and that for the listing's average shill scores is plotted in Figure 2 (the upper panel is for all listings, and the lower one is for listings receiving at least one bid). Note that although about 27% (2,954) of the auctions had a shill score of 0, this is not because bidding behavior in these auctions clearly indicated that they were free of shills, but simply because an overwhelming majority of them received no bid at all (2,757, see Table 6). If we consider only listings that received at least one bid, the percentage of auctions having a shill score of 0 will fall to 188, about 2.3% of the listings with at least one bid. We computed the correlation between the average shill scores and the reserve prices, which was -0.72 for all listings and -0.48 for listings receiving at least one bid. This is an important fact which confirms the intuition in the literature (e.g., Steigltitz, 2007; Kauffman and Wood, 2003) that the sellers who practice shill bidding often start with a low reserve price. Obviously, they

[16] The central idea of PCA is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible the variation present in the data. See, e.g., Jolliffe (2002).

start with a low reserve price to attract more bidders, and then compete with them to drive up the price.

3.2 The Effects of Shill Bids

In this section, we investigate the effect of the reserve price on trade probability and transaction price, while taking the influence of shilling into consideration. Since there is a transaction price only when there is a trade, we adopt the standard Heckman two-stage procedure for the estimation. In the first-stage, a probit model estimates the factors that affect the transaction probability of a car. In the second stage, an OLS model is used to estimate the transaction price of the items that are sold.

The first stage estimation is as follows:

$$\begin{aligned} Sold_j = {} & \alpha_0 + \alpha_1 \times StBidR + \alpha_2 \times BIN_j + \alpha_3 \times StBidR \times BIN_j \\ & + \alpha_4 \times SRP_j + \alpha_5 \times Competitor_j \\ & + \alpha_6 \times ln(SellerReputation_{m(j)}) + \alpha_7 \times Warranty_j \\ & + \alpha_8 \times AMileage_j + \alpha_9 \times Seller\ is\ Dealer_{m(j)} + \alpha_{10} \times Seller's\ Experience_{m(j)} \\ & + \alpha_{11,d} \times Posted\ Duration_{d(j)} + \alpha_{12,f} \times Fuel\ Type_{f(j)} + \alpha_{13,l} \times Model_{l(j)} \\ & + \alpha_{14,c} \times Car\ Body\ Type_{c(j)} + \alpha_{15,v} \times Vehicle\ Condition_{v(j)} + \varepsilon_{1,j} \end{aligned} \tag{4}$$

In the equation, $Sold_j$ equals 1 if the item is sold, and 0 otherwise. $m(j)$ is seller m in listing j, $d(j)$ = 1, 2, 3 is a dummy representing the auction duration of 3, 5, and 7 days, respectively; f = 1, ..., 6 are dummies which denote the vehicle's fuel type; l = 1, ..., 19 are dummies which denote the 19 models of vehicles. c= 1, ..., 9, which denotes the car body type. w = 1, 2 denotes whether the vehicle title bears the designation "clear" or "salvage".[17]

As shown in Chen et al. (2017), the seller has different values for the optimal reserve price in the regular and buy-it-now auctions. In particular, the optimal reserve price is higher in the

[17] For information on the characteristics of cars above, see Table 4.

buy-it-now auction than in the regular auction. In order to control for this difference, we used a dummy for the buy-it-now auction, BIN_j, which equals 1 if the car was listed in a buy-it-now auction.

The variable $StBidR$ is the ratio of starting bid to the car's Blue Book value. Since eBay does not allow the seller to post an open reserve price, the open reserve price is in the form of the starting bid, which the sellers must specify in the regular auction and the buy-it-now auction. In our model, we view the reserve price as the starting bid chosen by the seller, and the coefficient of which is one of our main interests. $AMileage$ is the average mileage of the car (i.e., the car's mileage divided by its age). Although the Blue Book value is supposed to be a summary statistics of a car's characteristics that affect its price, it is not directly related to trade probability. A higher Blue Book price does not imply a lower trade probability: it only reflects better characteristics of the car. Rather, characteristics such fuel type, model, body type, and vehicle condition can better control for the bidder's preference that might affect the car's trade probability.[18] We therefore used these variables to control for the car's trade probability. Also, since these characteristics are also strongly related to the Blue Book value, the latter was not used as a control variable.

The seller's experience, reputation, and whether she is a dealer are all expected to influence trade probability, and were used as control variables. For the seller's experience, some studies used the seller's feedback or reputation as a proxy (Kauffman and Wood, 2006; Hu and Wang, 2010; Newberry, 2015), while others used the number of days since the seller joined the auction platform (Chen et al., 2013; Scott, Gregg, and Choi, 2015). We used how many transactions the seller had made within a year as a proxy for experience, because we believe that it is more relevant to the current transaction. The number of competitors, as the number of similar cars on eBay Motors during the time an item was listed, is expected to negatively influence trade probability,

[18] For example, when there are media mentions of fossil fuels' greenhouse effect, the demand for hybrid cars might increase, even though their prices are higher.

while warranty and posted duration are expected to have a positive effect. Secret reserve price is widely known to have a negative effect on trade probability (see, for example, Katkar and Reiley, 2006; Bajari and Hortaçsu, 2003), and we used a dummy, SRP, to control for it. This variable is important, as in our sample more sellers used it than not (see Table 6).

Shill score was not used as a control variable in the trade equation, and there were two reasons for this. First, the reason the seller places shill bids is to influence the transaction price. Consequently, there is no reason for the seller to place a bid before there is any eligible bid, or when the bids are lower than the reserve price, secret or open. In other words, sellers place shill bids if and only if the item is sure to be sold. Therefore, shill bid should not affect the sale probability. Second, in our data 2,757 listings (regular auction: 1,144; buy-it-now auction: 1,613) had the lowest possible shill score (0) not because bidding behavior in these listings strongly suggested that they were not shilled, but simply because these listings received no bid. If we included shill score in the trade equation, since many listings received no bid (and were not sold) and had a shill score of 0, it would create an artificial positive relationship between shill score and trade probability that is not caused by shilling.

After estimating the sales equation, we then estimated the transaction price for the items that were sold. The second stage is an OLS estimation:

$$
\begin{aligned}
WinBidR; = {} & \beta_0 + \beta_1 \times StBidR + \beta_2 \times BIN_j + \beta_3 \times StBidR \times BIN_j \\
& + \beta_4 \times Average\ Shill\ Score_j \\
& + \beta_5 \times Blue\ Book\ Value_j + \beta_6 \times Competitor_j \\
& + \beta_7 \times ln(Seller\ Reputation_{m(j)}) + \beta_8 \times Warranty_j + \beta_9 \times AMileage_j \\
& + \beta_{10} \times Seller's\ Experience_{m(j)} + \beta_{11,d} \times Posted\ Duration_{d(j)} + \varepsilon_{2,j}.
\end{aligned} \tag{5}
$$

The dependent variable, $WinBidR$, is the winning bid divided by the car's Blue Book value. The variable $StBidR$ is the starting bid divided by the car's Blue Book value, whose coefficient is one of our main interests. The Blue Book value was used to control for the characteristics and type/brand of a car that influence its price. Note that although its value is expected to have a positive effect on price, we had already divided transaction price by its Blue Book value. Therefore, its effect on transaction price might not be significant. Seller's experience and reputation are both expected to have positive effects on the transaction price. The influence of warranty on car price is obvious.

The literature has adopted two measures when estimating how characteristics of used cars affect their price premiums. The first is the difference between the car's Blue Book value and its price. In this vein, Andrews and Benzing (2007) use two different measures for the premium/discount for each vehicle. The first is the difference between the highest bid of each listing and the Blue Book value of an automobile; and the second is the difference between the winning price and the Blue Book value. Both measures are based on price difference. The second is the ratio of the winning price to the Blue Book value instead of their differences, which was also adopted by Bajari and Hortaçsu (2003). This is also the measure that we adopted here.

We also followed Lucking-Reiley et al. (2007) to include dummy variables for various lengths of posted duration. The posted duration of 10 days was used as the basis of comparison. Moreover, as indicated in Newberry (2015), the winning price of a car can be influenced by how many similar cars are in competition with each other, and by the mileage and age of the car. Therefore, we included two variables, $Competitor$ and $AMileage$, to control for the effects of competition and vehicle's average mileage.

Despite the fact that shill bidding is a well-recorded behavior in online auctions, it is not prevalent after all. For example, Kauffman and Wood (2003) found that the proportion of bids that resembled shills among all bids was 2.84%.[19] In Engelberg and Williams (2009), the proportion was 1.39%. The proportions of bids that are highly likely to be shills are small in both studies.[20]

[19] There were 30,496 bids in their data, among which 866 were deemed questionable.

[20] Note that the two percentages are almost surely under-estimated, because they only account for those who are *highly likely to be shill bidders.*

However, in our calculation, 1,935 listings (17.8%; and is 23.8% among listings that receive at least one bid) have a shill score of at least 0.5. This raises a concern for our study that, since we used shill score as a continuous variable, we might have overestimated the effect of shill bids for listings that were actually not shilled, and underestimated those that were.

If we assume that the proportion of shill bids among all bids is about the same as the proportion of listings that are shilled among all listings, then since there are 10,893 listings in our sample, the number of listings that were shilled is 151 if the proportion is 1.39%, and is 309 if 2.84%. Therefore, we constructed a dummy which equals 1 if a listing's shill score is in the top 309, and 0 if not.[21] We then run an alternative regression which was identical to equations (4) and (5), except that we used the shill dummy as the control variable, rather than the shill score.

4 Results and Discussion

Table 7 presents the estimation results when using the shill score as a control variable, and Table 8 presents the results using the shill dummy. The second and third columns are the regression results without controlling for the shill bid, while the last two columns are the results that control for it. Except for the coefficient of BIN, which becomes significant, Table 8 is qualitatively identical to Table 7. In other words, using the shill dummy only quantitatively, not qualitatively, changes the regression results. The Variance Inflation Factors for the second stage were calculated to check collinearity between the starting price and the shill score, but we did not find serious collinearity in the model.

In both tables, the reserve price's effect on transaction probability is negative and highly significant regardless of whether shill score is controlled for (both are -1.66, and in both tables), a result not only consistent with almost all literature,[22] but (since the coefficients are almost

[21] We have also used 151 as the criterion, and the qualitative results were identical.

[22] Ariely and Simonson (2003), Bajari and Hortaçsu (2003), Häubl and Popkowski Leszezye (2003), Reiley (2006), Ku et al. (2006), Simonsohn and Ariely (2008), Brown and Morgan (2009), Barrymore and Raviv (2009),

identical) also confirms the intuition that phantom bids are not meant to increase trade probability. The effect of reserve price on the transaction price is positive and highly significant, again regardless of whether the shill score is controlled for. This is consistent with the literature which shows a positive effect of reserve price on price,[23] but inconsistent with some others.[24] However, this effect is more pronounced when shill bidding is controlled for, a result that is quite intuitive. When we do not control for the shill bids, the effect of reserve price on the transaction price is under-estimated, because the reduction in transaction price, when the seller reduces the reserve price, has been recovered by the seller's phantom bids. When we filter the effect of shill bids, its magnitude becomes what it should have been. Indeed, as can be seen from Table 7, the impact of reserve price on the transaction price is more than doubled, from 0.182 to 0.386, when we use shill score to control for shill bids' effect.

The coefficient for shill score in Table 7 is 0.56 and is significant at the 1% level, implying that a 0.1 increase in shill score will increase the ratio of transaction price to Blue Book price by about 0.056. Therefore, as long as the listing's shill score is greater than $0.386/0.556 \simeq 0.69$, then reducing the reserve price actually increases the transaction price. In other words, in the listings which are highly likely to be shilled, the sellers actually increase the transaction price by setting a low reserve price. This is consistent with the literature (Steiglitz, 2007; Kauffman and Wood, 2003) which postulates that the sellers who practice shill bids generally set a lower reserve price, then drive up the transaction price by placing phantom bids, to a level that is even higher than one with a reserve price consistent with theory. In this sense, that auctions with a low reserve price often yield higher revenue might be an illusion caused by shill bidding. Note that although the reserve price's effect on the transaction price has a steeper slope when shill bid is controlled

Einav et al. (2015), and Choi et al (2016).

[23] Ariely and Simonson (2003), Häubl and Popkowski Leszezye (2003), Reiley (2006), Brown and Morgan (2009), Barrymore and Raviv (2009), Einav et al. (2015), and Choi et al (2016)

[24] Bajari and Hortaçsu (2003), Kamins et al. (2004), Ku et al. (2006), and Simonsohn and Ariely (2008) find a negative effect on transaction price; Lucking-Reiley et al. (2007) and Einav et al. (2015) find no effect of reserve price.

for, it also has a smaller intercept. This is exactly because the sellers who shill usually start with a low price.

If we look at the shill dummy, rather than the shill score, Table 8 shows that the shill dummy increases the winning bid ratio by 0.11. That is, shill bidding increases the ratio of transaction price to Blue Book price by 11%. For example, if the Blue Book price of a car is $2,000, and assuming that the transaction price is usually 70% of the Blue Book price,[25] then the seller who shills will increase the transaction price by about $154. Also note that the effects of the reserve price on the transaction probability and the transaction price in Table 8 are almost the same between when shill bid is controlled and not (-1.66 for the trade probability, and -0.18 for transaction price). Therefore, these can be seen as the "true" effects of the reserve price without shill bidding.

Most of the other variables also have influences on transaction probability and transaction price that are consistent with intuition and the literature. For example, longer listing duration tends to result in higher transaction probability and transaction price.[26] Higher mileage reduces transaction price, and warranty increases transaction price. Warranty also has a negative effect on transaction probability, perhaps exactly because cars with warranty are more expensive. Dealers have a harder sale, a result that is also found in several other studies (e.g. Andrews and Benzing, 2007; Lewis, 2011). Secret reserve price reduces probability of transaction, which is consistent with the literature.[27] Seller's reputation increases sales and price, which is quite obvious and is well recorded.[28] Seller's experience negatively affects transaction price, but the coefficient is very small. The number of simultaneous competing listings reduces transaction probability, but increases transaction price. The former result is intuitive, but we cannot explain the latter.

[25] In our data, the mean value of $WinBidR$ is 0.7.

[26] In the regression, a 10-day duration is the base of comparison.

[27] Katkar and Reiley (2006), and Bajari and Hortaçsu (2003).

[28] Houser and Wooders (2006), Livingston (2005), Bajari and Hortaçsu (2003), and McDonald and Slawson (2002).

Finally, the Blue Book value does not affect transaction price. As explained earlier, this is because the transaction price has already been divided by the Blue Book value.

The buy-it-now dummy has a negative coefficient, and has twice as great a magnitude as that for the cross term of reserve price and the buy-it-now dummy. Since the value of *StBidR* is almost always smaller than 1, this implies that the buy-it-now auction has a smaller transaction probability than the regular auction. This, however, is not a general result, but comes from the fact that our sample contains only those buy-it-now listings which were not sold by buy-it-now price. Since we excluded the listings which were sold with buy-it-now prices from our sample, the excluded buy-it-now listings were all sold. This naturally reduces the sale probability of the buy-it-now items. Moreover, since the buy-it-now and reserve price cross term has a positive coefficient, it also implies that the reserve price has smaller negative effect on transaction probability for the buy-it-now listings. This is a reasonable result. As shown in Chen et al. (2017), the seller usually sets a higher reserve price in a buy-it-now auction than in a regular auction. Since there is no reason that a buy-it-now auction will have a higher trade probability when the seller shills (note that a buy-it-now auction has already reduced into a regular auction in our sample), the smaller coefficient for the buy-it-now auction simply adjusts for the effect of its higher reserve price.

For the effect on expected revenue, we first define expected revenue as:

$$Expected\ Revenue = Pr(Sold_j \mid StBidR) \times (WinBidR \mid StBidR). \tag{6}$$

The trade probability and the winning price of the vehicle are both estimated by substituting *StBidR* of the listings into the first- and the second-stage models, while for the value of other variables we take the sample mean. Figure 3 then plots with expected revenue calculated under the two specifications in Table 7. As shown in Figure 3, reserve price has an inverse relationship to expected revenue, regardless of whether shill bidding is controlled for. Therefore, as least for our data, low reserve price remains a puzzlingly good strategy if the seller aims to maximize expected revenue. However, the expected revenue is flatter when controlling for shill bid than when not. This means after controlling for shill bids, the benefit of setting a lower starting price drops, suggesting that part of such a benefit should be attributed to shill bidding. As can be seen from the upper two figures, the benefit mainly comes from its effect on transaction price. What our result shows is that, when the seller sets a lower reserve price, although the seller's expected revenue does not increase as much as when she shill bids, it does still increase.

5 Last-Minute Shill Bids

In a recent paper, Bose and Daripa (2017) propose a new theory of sniping,[29] based on the bidder's strategic reaction to shill bids. When the bidders are aware of the possibility of shill bids, a strategic response for them to avoid competing with the shill bidders is to delay placing bid and, in particular, to place bids right before the last minute of the auction. In that case, the only possible time for the sellers to place phantom bids is during the last minute, when there is a positive probability that their bids cannot get through.

If this theory is true, then shill bids will occur primarily before the end of the auctions. The empirical implication for this is that there will be a higher percentage of shill bids during the last minutes of an auction than before the last minute. It should, however, be emphasized that this theoretical prediction actually conflicts with the common thinking in the literature. As can be seen in Section 3.1, most of the literature takes early bids, rather than last-minute bids, as one of the signs of shill bids.[30] In fact, in defining shill score in equation (1), this criterion is explicitly taken into consideration.[31] To do justice to the theory, we first redefine shill score by deleting the last two criteria (i.e., deleting $\zeta_{i,j}$ and $\eta_{i,j}$) in its calculation, then recalculate the remaining five

[29] Sniping refers to the bidder's oft-observed tendency in online auctions of withholding their bids until the last moment before the auction closes. Therefore, a stylized fact in the online auctions is that very few bids are placed long after the auction begins, but a bidding war starts right before it ends. See Steiglitz (2007) and Ockenfels et al. (2006).

[30] Kauffman and Wood (2003); Xu et al. (2009); Dong et al. (2009); Trevathan and Read (2009), and Liu (2017).

[31] See characteristic j_i before the definition of shill score in Section 3.1.

coefficients (θ_i, i=1,...,5) by PCA. Finally, we propose three definitions on what constitutes the "last minute" of an auction.

The coefficients for the shill score resulting from PCA change substantially after $\zeta_{i,j}$ and $\eta_{i,j}$ are deleted. The values of θ_i's become $\theta_1 = 1$, $\theta_2 = 2$, $\theta_3 = 7$, $\theta_4 = 5$, and $\theta_5 = 5$. That is, the shill score for bidder i and listing j becomes

$$Shill\ Score_{i,j} = \frac{\alpha_{i,j} + 2\beta_{i,j} + 7\gamma_{i,j} + 5\delta_{i,j} + 5\varepsilon_{i,j}}{20}.$$

Similar to what constitutes shill bidding, there is not a precise mathematical definition of sniping, and especially of what constitutes the "last minute". Researchers usually use the last 5 or 10 minutes of an auction as a threshold.[32] We use the last 30 minutes, the last 5 minutes, and the last 5% of the auction duration as three possible thresholds for the last minute. Our aim is to test whether the average value of the shill scores for the bidders who place bids after the threshold is greater than those before. If the prediction in Bose and Daripa (2017) is correct, then the former should be substantially and significantly greater. The data we use is the sample for the first part of our empirical study. Using the above-mentioned three thresholds, we compare the average value of the bidders' shill scores who bid before and after the thresholds. The results are summarized in Table 9. As can be seen from the table, contrary to the theoretical prediction, the bidders' average shill scores are consistently higher for those who bid before, rather than after, the thresholds for all three measures. The t-test shows that the differences are all significant at the 1% level.

An alternative test for the prediction is to see whether the bidders who are most likely to be the shill bidders predominantly bid in the last minute. Following Section 4, we take the bidders whose shill scores are at the top 2.84% as the shill bidders, and calculate the proportions that these bidders constitute among all bidders during the last minute and before the last minute. The

[32] See the survey in Ockenfels et al. (2006).

results are reported in Table 10. Similar to the results in Table 9, the proportion of shill bidders who appear during the last minute is smaller, which is significant at the 1

6 Conclusion

In this paper, we reinvestigate the relationship between the reserve price and the auction outcomes. We consider and control for a factor that might affect this relationship, but has so far escaped researchers' attention: the seller's shill-bid behavior. By constructing the shill score for each listing and controlling for its impact in the regressions, we show that shill bids do not affect the sale rate but increase the transaction price conditional on sale. Moreover, an increase in reserve price reduces trade probability but increases the transaction price. Therefore, for items with low reserve prices, the find transaction prices are partly driven up by shill bidding. However, even after controlling for its influence, lower reserve price still results in higher expected revenue for the seller. Given that participating in an online auction incurs low cost, and given that the private-valuation model seems a good fit for bidders' willingness-to-pay for used cars, it is unlikely that either the entry cost or the correlation-value explanation mentioned in the Introduction is the reason behind our results. As such, unless we believe that the power of low reserve price in creating bidding fever is so strong and prevalent, otherwise the low reserve price - high revenue relationship remains a mystery. Finally, we show that bidders who snipe have lower average shill scores, and they place bids earlier than the average bidders. These results contradict the theoretical prediction of Bose and Daripa (2017). Possible explanations might be that the bidders are not aware of the possibility of shills when they bid, or do not feel they warrant attention, or the sellers might think that placing shill bids in the last minute is too risky.

References

[1] Adam, M.T.P, Krämer, J., Jähnig, C., Seifert, S., and Weinhardt, C., (2011), "Understanding Auction Fever: A Framework for Emotional Bidding," *Electronic Markets*, 21, 197–207.

[2] Adam, M.T.P., Krämer, J., and Müller, M.B. (2015), "Auction Fever! How Time Pressure and Social Competition Affect Bidders'Arousal and Bids in Retail Auctions," *Journal of Retailing*, 91(3), 468-485.

[3] Andrews, T., and Benzing, C. (2007), "The Determinants of Price in Internet Auctions of Used Cars," *Atlantic Economic Journal*, 35(1), 43-57.

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

[5] Bag, P.K., Dinlersoz, E.M., Wang, R. (2000), "More on Phantom Bidding," *Economic Theory*, 15(3), 701–707.

[6] Bajari, P., and Hortaçsu, A. (2003), "The Winner's Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions," *The RAND Journal of Economics*, 34(2), 329-355.

[7] Barrymore, N., and Raviv, Y. (2009), "The Effect of Different Reserve Prices on Auction Outcomes," Robert Day School of Economics and Fincance Research Working Paper No. 2009-13.

[8] Bose, S., and Daripa, A. (2017), "Shills and Snipes," *Games and Economic Behavior*, 104, 507-516.

[9] Brown, J., and Morgan, J. (2009), "How Much Is a Dollar Worth? Tipping Versus Equilibrium Coexistence on Competing Online Auction Sites," *Journal of Political Economy*, 117(4), 668-700.

[10] Chakraborty, I. and Kosmopoulou, G. (2004). "Auctions with Shill Bidding," *Economic Theory*, 24(2), 271-282.

[11] Chen, J.R., Chen, K.P., Chou, C.F. and Huang, C.I. (2013), "A Dynamic Model of Auctions with Buy-it-Now: Theory and evidence," *Journal of Industrial Economics*, 61(2), 393-429.

[12] Chen, K.P., Ho, S.H., Liu, C.H. and Wang, C.M. (2017), "Optimal Listing Strategy in Online Auctions," *International Economic Review*, 58(2), 421-437.

[13] Choi, S., Nesheim, L., and Rasul, I. (2016), "Reserve Price Effects in Auctions: Estimates from Multiple Regression-Discontinuity Designs," *Economic Inquiry*, 54(1), 294-314.

[14] Dong, F., Shatz, S.M., and Xu, H. (2009), "Inference of Online Auction Shills using Dempster-Shafer Theory," *2009 Sixth International Conference on Information Technology: New Generations* (pp. 908-914). Washington DC: IEEE Computer Society.

[15] Engelberg, J. and Williams, J. (2009). "eBay's Proxy Bidding: A License to Shill," *Journal of Economic Behavior and Organization*, 72(1), 509 - 526.

[16] Engelbrecht-Wiggans, R., and Nonnenmacher, T. (1999), "A Theoretical Basis for 19th-Century Changes to the Port of New York Imported Goods Auction," *Explorations in Economic History*, 36, 232-245.

[17] Einav, L., Kuchler, T., Levin, J., and Sundaresan, N. (2015), "Assessing Sale Strategies in Online Markets Using Matched Listings," *American Economic Journal: Microeconomics*, 7(2), 215-247.

[18] Graham, D.A., Marshall, R.C., Richard, J.-F. (1990), "Phantom Bidding against Heterogenous Bidders," *Economics Letters*, 32, 13–17.

[19] Häubl, G., and Popkowski Leszczyc, P. (2003), "Minimum Prices and Product Valuations in Auctions," *Marketing Science Institute Reports*, 03-117.

[20] Izmalkov, S. (2004), "Shill Bidding and Optimal Auctions," MIT Working Paper.

[21] Jolliffe, I. T. (2002), "Principal Component Analysis and Factor Analysis," *Principal Component Analysis*, New York, NY, Springer.

[22] Jones, M.T., (2011), "Bidding Fever in eBay Auctions of Amazon.com Gift Certificates," *Economics Letters*, 113, 5-7.

[23] Kamins, M.A. , Dreze, X. , Folkes, V.S. (2004), "Effects of Seller-supplied Prices on Buyers? Product Evaluations: Reference Prices in an Internet Auction Context," *Journal of Consumer Research* 30, 622–628 .

[24] Katkar, R. , Reiley, D.H. (2006), "Public versus Secret Reserve Prices in eBay Auctions: Results from a Pokémon Field Experiment," *Advances in Economic Analysis and Policy* 6 (2), 7.

[25] Kauffman, R.J., and Wood, C.A. (2003), "Running up the Bid: Detecting, Predicting, and Preventing Reserve Price Shilling in Online Auctions," *ICEC*. doi: 10.1145/948005.948040.

[26] Kauffman, R. J., and Wood, C. A. (2005), "The Effects of Shilling on Final Bid Prices in Online Auctions," *Electronic Commerce Research and Applications*, 4(1), 21-34.

[27] Kauffman, R.J. and Wood, C.A. (2006), "Doing their Bidding: An Empirical Examination of Factors that Affect a Buyer's Utility in Internet Auctions," *Information Technology and Management*, 7(3), 171-190.

[28] Kosmopoulou, G. and De Silva, D. G. (2007). "The Effect of Shill Bidding upon Prices: Experimental Evidence," *International Journal of Industrial Organization*, 25(2), 291-313.

[29] Ku, G., Galinsky, A.D., and Murnighan, J.K. (2006), "Starting Low but Ending High: A Reversal of the Anchoring Effect in Auctions," *Journal of Personality and Social Psychology*, 90(6), 975-986.

[30] Ku, G., Malhotra,D., and Murnighan, J.K. (2005), "Towards a Competitive Arousal Model of Decision-Making: A Study of Auction Fever in Live and Internet Auction," *Organizational Behavior and Human Decision Processes*, 96, 89-103.

[31] Heyman, J.E., Orhun, Y., and Ariely, D. (2004), "Auction Fever: The Effect of Opponents and Quasi-Endowment on Product Valuations,"*Journal of Interactive Marketing*, 18, 7-21.

[32] Houser, D., and Wooders, J. (2006), "Reputation in Auctions: Theory and Evidence from eBay," *Journal of Economics and Management Strategy*, 15, 353–369.

[33] Hu, Y., and Wang, X. (2010), "Country-of-Origin Premiums for Retailers in International Trades: Evidence from eBay's International Markets," *Journal of Retailing*, 86(2), 200-207.

[34] Levin, D., and Smith, J.L. (1994), "Equilibrium in Auctions with Entry," *American Economic Review*, 84, 585-599.

[35] Levin, D., and Smith, J.L. (1996), "Optimal Reservation Prices in Auctions," *Economic Journal*, 106, 1271-1283.

[36] Lewis, G. (2011), "Asymmetric Information, Adverse Selection and Online Disclosure: The Case of eBay Motors," *American Economic Review*, 101, 1535–1546.

[37] Liu, Y.C. (2017), "Applying Clustering to Analyze Bidding Behaviors and Shill Bidding in Online Auction," Master's Dissertation. National Sun Yat-Sen University, Taiwan.

[38] Livingston, J. A. (2005), “How valuable is a good reputation? A Sample Selection Model of Internet Auctions,” *The Review of Economics and Statistics*, 87(3), 453–465

[39] Lucking-Reiley, D., Bryan, D., Prasad, N. and Reeves, D. (2007), “Pennies from eBay: The Determinants of Price in Online Auctions,” *The Journal of Industrial Economics*, 55(2), 223-233.

[40] Mamun, K. (2015), “Combating Shill Bidding in Real Time: Prevention, Detection and Response,” *Computer and Information Science*, 8(2), 24-36.

[41] Mamun, K., and Sadaoui, S. (2013), “Combating Shill Bidding in Online Auctions,” *International Conference on Information Society, i-Society.*, 170-176.

[42] McAfee, R. P., and McMillan, J. (1987). “Auctions and Bidding,” *Journal of Economic Literature*, 25(2), 699-738.

[43] McDonald, C.G. , Slawson, V.C. (2002), “Reputation in an Internet Auction Market,” *Economic Inquiry*, 40, 633–650 .

[44] McGrath, S. and McGrath, L. (2010), “The Complete Idiot’s Guide to Ebay,” 2nd Edition. Indianapolis, Indiana, USA: Alpha Books.

[45] Milgrom, P., and Weber, R. (1982), “A Theory of Auctions and Competitive Bidding,” *Econometrica*, 50(5), 1089-1122.

[46] Myerson, R. B. (1981). “Auction Design,” *Mathematics of Operations Research*, 6(1), 58-73.

[47] Newberry, P.W. (2015), “The Effect of Competition on eBay,” *International Journal of Industrial Organization*, 40, 107-118.

[48] Ockenfels, Reiley, D., and Sadrieh, A. (2006), "Online Auctions," In T.J. Hendershott (Ed.), Handbooks in Information Systems I, Handbook on Economics and Information Systems, pp.571-628, Amsterdam: Elsevier Science.

[49] Riley, J. G. and Samuelson, W. F. (1981). "Optimal Auctions," *The American Economic Review*, 71(3), 381-392.

[50] Reiley, D. (2006),"Field Experiments on the Effects of Reserve Prices in Auctions: More Magic on the Internet," RAND *Journal of Economics* 37(1), 195–211 .

[51] Scott, J.E., Gregg, D.G. and Choi, J.H. (2015), "Lemon Complaints: When Online Auctions Go Sour," *Information Systems Frontier*, 17, 177-191.

[52] Shah, H.S., Joshi, N.R., Sureka, A., and Wurman P.R. (2003), "Mining eBay: Bidding Strategies and Shill Detection," In: Zaïane O.R., Srivastava, J., Spiliopoulou, M., Masand, B. (eds). WEBKDD 2002 - Mining Web Data for Discovering Usage Patterns and Profiles. *WebKDD 2002. Lecture Notes in Computer Science, 2703.* Berlin, Heidelberg, Germany: Springer.

[53] Simonsohn, U., and D. Ariely. (2008), "When Rational Sellers Face Non-Rational Consumers: Evidence from Herding on eBay," *Management Science*, 54(9), 1624–37.

[54] Steiglitz, K. (2007), *Snipers, Shills, and Sharks: eBay and Human Behavior*, Princeton University Press, Princeton.

[55] Trevathan, J., and Read, W. (2009), "Detecting Shill Bidding in Online English Auctions," *Handbook of Research on Social and Organizational Liabilities in Information Security*, 46, 446-472.

[56] Vincent, Daniel, (1995), "Bidding Off the Wall: why Reserve Prices May be Kept Secret?" *Journal of Economic Theory*, 65, 575–558.

[57] Xu, H, Bates, C.K., and Shatz, S.M. (2009), "Real-Time Model Checking for Shill Detection in Live Online Auctions," *Software Engineering Research and Practice*, 134-140.

Table 1: The Definition and Description of Variables for Identifying IDs and Constructing Shill Scores

Variables	Description
Buyer Characteristic	
$NoBids_k$	The number of bids for bidder i participated in listing j.
$RepB_k$	The bidder's reputation for the kth bid in listing j.
$BidAmout_k$	The amount of the bidder i's kth bid in listing j.
$BidIncrement_k$	The amount difference between bids placed by bidder i and the latest bid before that in listing j.
$InterBidTime_k$	The time difference between bids placed by bidder i and the latest bid before that in listing j.
$DiffFirstBid_k$	The difference between the expiration time of listing j and the time of bidder i's first bid.
$DiffLastBid_k$	The difference between the expiration time of listing j and the time of bidder i's last bid.
Seller Characteristic	
n	The number of listings held by seller $m(j)$.
Shill Indexs	
$\alpha_{i,j}$	The percentage of the seller's auctions bidder i has participated in given a particular seller.
$\beta_{i,j}$	The percentage of bids that bidder i has submitted in listing j.
$\gamma_{i,j}$	The proportion of losses that bidder i has participated in given a particular seller.
$\delta_{i,j}$	The normalized average inter-bid times for bidder i participated. in listing j.
$\varepsilon_{i,j}$	The normalized average inter-bid increments for bidder i participated in listing j.
$\zeta_{i,j}$	The normalized time differences between the expiration time and the time of bidder i's first bid in listing j.
$\eta_{i,j}$	The normalized time differences between the expiration time and the time of bidder i's last bid in listing j.
$ShillScore_{i,j}$	The shill score for bidder i in listing j.

Table 2: **The Summary Statistics for Identifying IDs and Constructing Shill Scores**

Panel A: Basic Information		**All**	**RA**	**BINA**
The number of listings		18,441	9,473	8,968
The number of bids		154,599	94,082	60,517
for bidder's ID partially concealed		151,230	91,845	59,385
for bidder's ID fully revealed		3,369	2,237	1,132
The number of sellers		7,653	-	-

Panel B: Buyer and Seller Characteristics				
Bid level (154,599 obs.)	**Mean**	**S.D.**	**Min**	**Max**
$NoBids$	4.44	5.21	1.00	67.00
$RepB$	89.34	480.72	0.00	70014.00
$BidAmount\ (USD)$	5842.59	6050.12	0.01	67775.00
$BidIncrement\ (USD)$	369.91	744.82	0.00	39300.00
$InterBidTime\ (second)$	28298.67	52532.39	0.00	851640.00
$DiffFirstBid\ (day)$	3.45	2.66	0.00	11.00
$DiffLastBid\ (day)$	2.92	2.64	0.00	11.00
Seller-level (7,653 obs.)	**Mean**	**S.D.**	**Min**	**Max**
n	2.41	6.39	1.00	281.00

Panel C: For Constructing Shill Scores				
Bidder Listing level (72,597 obs.)	**Mean**	**S.D.**	**Min**	**Max**
α	0.10	0.14	0.00	1.00
β	0.29	0.27	0.00	1.00
γ	0.26	0.36	0.00	0.98
δ	0.60	0.42	0.00	1.00
ε	0.61	0.38	0.00	1.00
ζ	0.49	0.42	0.00	1.00
η	0.47	0.42	0.00	1.00
$ShillScore$	0.43	0.21	0.00	1.00

Note: RA represents regular auction, and BINA represents buy-it-now auction.

Table 3: The Definition and Description of Variables for Regular Auctions

Variables	Description
Transaction information	
$Sold_j$	A dummy variable indicating whether vehicle in listing j is sold or not.
$WinBidR$	The winning price of listing j divided by listing j's Kelley Blue Book value.
Auction characteristic	
$StBidR$	The starting price of listing j divided by listing j's Kelley Blue Book value.
BIN_j	A dummy variable indicating whether the listing j is listed under buy-it-now auction.
SRP_j	A dummy variable indicating whether the listing j has set a secret reserve price.
$Posted\ Duration_{d(j)}$	Dummy variables indicating whether the duration of the listing j is 3 days, 5 days, 7 days, or 10 days.
$Average\ Shill\ Score_j$	The average shill score of listing j.
$Blue\ Book\ Value_j$	The Kelley Blue Book value for the car in listing j. (In 1,000 USD)
$Competitor_j$	The number of vehicles with the same model and age as listing j listed auction within the posted duration of j.
Seller characteristic	
$ln(Seller\ Reputation_{m(j)})$	The natural log of seller $m(j)$'s reputation.
$Seller\ is\ Dealer_{m(j)}$	A dummy variable indicating whether the seller $m(j)$ is a car dealer.
$Seller's\ Experience_{m(j)}$	How many transactions the seller $m(j)$ has made within a year.
Car characteristic	
$Warranty_j$	A dummy variable indicating whether the vehicle in listing j has warranty or not.
$AMileage_j$	The mileage of the vehicle in listing j divided by its age. (In 1,000 mile/year)
$Car\ Model_j$	The car model of the vehicle in listing j. There are 20 car models in our sample.
$Vehicle\ Condition_j$	Whether the vehicle condition is Clear, Salvage, or Other in listing j.
$Car\ Body\ Type_j$	The car model of the vehicle in listing j. There are 10 body types in our sample.
$Fuel\ Type_j$	The fuel type of the vehicle in listing j. There are 7 fuel types in our sample.

Table 4: **The Summary Statistics of Auction Variables**

Variable	**Obs.**	**Mean**	**S.D.**	**Min**	**Max**
Trade Information					
$Sold_j$	10,893	0.171	0.376	0	1
$WinBidR$	1,858	0.696	0.281	0.046	2.188
Auction characteristic					
$StBidR$	10,893	0.392	0.393	0.00000028	2.778
BIN_j	10,893	0.516	0.500	0	1
SRP_j	10,893	0.765	0.424	0	1
$Posted\ Duration_{d(j)} = 3Days$	10,893	0.041	0.198	0	1
$Posted\ Duration_{d(j)} = 5Days$	10,893	0.136	0.343	0	1
$Posted\ Duration_{d(j)} = 7Days$	10,893	0.668	0.471	0	1
$Posted\ Duration_{d(j)} = 10Days$	10,893	0.155	0.362	0	1
$Average\ Shill\ Score_j$	10,893	0.287	0.210	0	0.964
$Blue\ Book\ Value_j$	10,893	14.188	6.502	1.575	47
$Competitor_j$	10,893	17.046	15.846	0	109
Seller characteristic					
$ln(Seller\ Reputation_{m(j)})$	10,893	4.181	1.780	0	9.571
$Seller's\ Experience_{m(j)}$	10,893	60.614	203.552	0	8372
$Seller\ is\ Dealer_{m(j)}$	10,893	0.171	0.450	0	1
Car characteristic					
$Warranty_j$	10,893	0.362	0.481	0	1
$AMileage_j$	10,893	15.067	15.844	0	999.999

Table 4: The Summary Statistics of Auction Variables (Continue)

Variable	Obs.	Mean	S.D.	Min	Max
Car characteristic					
Car Model					
4Runner	10,893	0.101	0.302	0	1
Avalon	10,893	0.035	0.183	0	1
Camry	10,893	0.150	0.357	0	1
Celica	10,893	0.029	0.168	0	1
Corolla	10,893	0.079	0.270	0	1
FJ Cruiser	10,893	0.026	0.159	0	1
Highlander	10,893	0.052	0.221	0	1
Land Cruiser	10,893	0.029	0.167	0	1
MR2	10,893	0.009	0.095	0	1
Matrix	10,893	0.019	0.136	0	1
Prius	10,893	0.068	0.253	0	1
RAV4	10,893	0.037	0.188	0	1
Sequoia	10,893	0.047	0.211	0	1
Sienna	10,893	0.058	0.235	0	1
Solara	10,893	0.043	0.204	0	1
Supra	10,893	0.012	0.107	0	1
Tacoma	10,893	0.123	0.329	0	1
Tercel	10,893	0.003	0.056	0	1
Tundra	10,893	0.067	0.250	0	1
Yaris	10,893	0.012	0.111	0	1

Table 4: **The Summary Statistics of Auction Variables (Continue)**

Variable	Obs.	Mean	S.D.	Min	Max
Car characteristic					
Vehicle Condition					
Clear	10,893	0.933	0.250	0	1
Salvage	10,893	0.051	0.221	0	1
Other	10,893	0.015	0.123	0	1
Car Body Type					
Convertible	10,893	0.032	0.176	0	1
Coupe	10,893	0.046	0.210	0	1
Hatchback	10,893	0.068	0.251	0	1
Minivan/Van	10,893	0.055	0.228	0	1
Pickup truck	10,893	0.185	0.388	0	1
SUV	10,893	0.282	0.450	0	1
Sedan	10,893	0.290	0.454	0	1
Wagon	10,893	0.011	0.106	0	1
Other	10,893	0.004	0.064	0	1
Unspecified	10,893	0.026	0.160	0	1
Fuel Type					
CNG	10,893	0.000	0.010	0	1
Diesel	10,893	0.000	0.014	0	1
Electric	10,893	0.000	0.019	0	1
Gasoline	10,893	0.941	0.235	0	1
Hybrid − electric	10,893	0.051	0.221	0	1
Other	10,893	0.006	0.079	0	1
Unspecified	10,893	0.000	0.019	0	1

Table 5: The Number of Listings, Bids, Bidders, and Sellers

Panel A: The First Part Sample

	All
The number of listings	18,841
The number of bids	154,599
The number of bidders	52,685
The number of sellers	7,653

Panel B: The Second Part Sample

	RA	**BINA**	**All**
The number of listings	5,268	5,625	10,893
The number of bids			69,919
The number of bidders			25,896
The number of sellers			4,433

Table 6: Secret Reserve Price and the Number of Bids

	Listing with No Bids	Listing with at Least One Bid
Listing w/o *SRP*	1,294	1,268
Listing with *SRP*	1,463	6,868
Bid greater than *SRP*	-	676
Bid smaller than *SRP*	-	6,192
The number of listings	2,757	8,136

Table 7: Regression Result of Heckman Model: Shill Score

	W/O Shill Score Control		With Shill Score Control	
	First Stage	**Second Stage**	**First Stage**	**Second Stage**
Variables	*Sold*	*WinBidR*	*Sold*	*WinBidR*
StBidR	-1.661***	0.182***	-1.661***	0.386***
	(0.0714)	(0.0257)	(0.0711)	(0.0296)
BIN	-0.535***	-0.067**	-0.528***	-0.0140
	(0.0496)	(0.0262)	(0.0495)	(0.0255)
$StBidR \times BIN$	0.256**	0.026	0.245**	-0.0200
	(0.1010)	(0.0448)	(0.101)	(0.0433)
Average Shill Score				0.556***
				(0.0430)
Blue Book Value $(k\$)$		-0.000516		-0.000371
		(0.00117)		(0.00112)
SRP	-1.613***		-1.617***	
	(0.0464)		(0.0457)	
Competitor	-0.00957***	0.00234***	-0.00965***	0.00237***
	(0.00152)	(0.000575)	(0.00153)	(0.000549)
ln(Seller Reputation)	0.0236**	0.00972**	0.0238**	0.00721*
	(0.0115)	(0.00401)	(0.0116)	(0.00382)
Seller's Experience	5.80e-05	-3.42e-05*	5.76e-05	-2.65e-05*
	(0.000105)	(1.98e-05)	(0.000104)	(1.87e-05)
Seller is Dealer	-0.322***		-0.351***	
	(0.0406)		(0.0403)	
Warranty	-0.142***	0.0968***	-0.140***	0.0970***
	(0.0447)	(0.0181)	(0.0448)	(0.0172)
AMileage	0.000913	-0.00101***	0.000936	-0.000932***
	(0.00919)	(0.000245)	(0.00896)	(0.000233)
$Duration = 3Days$	-0.0331	-0.0839***	-0.0300	-0.0754***
	(0.0939)	(0.0296)	(0.0941)	(0.0282)
$Duration = 5Days$	0.0743	-0.0466**	0.0802	-0.0321
	(0.0656)	(0.0222)	(0.0658)	(0.0212)
$Duration = 7Days$	-0.0075	-0.0158	-0.0748	-0.0154
	(0.0519)	(0.0183)	(0.0521)	(0.0175)
Constant	0.803**	0.511***	0.766*	0.279***
	(0.407)	(0.0310)	(0.413)	(0.0340)
Fuel Type	Y	N	Y	N
Car Model	Y	N	Y	N
Car Body Type	Y	N	Y	N
Vehicle Condition	Y	N	Y	N
ath Rho		0.482***		0.357***
		(0.0751)		(0.0728)
ln(Sigma)		-1.334***		-1.397***
		(0.0211)		(0.0196)
Observations	10,893	10,893	10,893	10,893

Table 8: **Regression Result of Heckman Model: Shill Dummy**

	W/O Shill Dummy Control		With Shill Dummy Control	
	First Stage	**Second Stage**	**First Stage**	**Second Stage**
Variables	*Sold*	$WinBidR$	*Sold*	$WinBidR$
$StBidR$	-1.661***	0.182***	-1.660***	0.188***
	(0.0714)	(0.0257)	(0.0714)	(0.0256)
BIN	-0.535***	-0.0668**	-0.537***	-0.0666**
	(0.0496)	(0.0262)	(0.0497)	(0.0260)
$StBidR \times BIN$	0.256**	0.0258	0.258**	-0.0222
	(0.1010)	(0.0448)	(0.102)	(0.0446)
Shill Dummy				0.110***
				(0.0291)
$Blue\ Book\ Value\,(k\$)$		-0.000516		-0.000471
		(0.00117)		(0.00116)
SRP	-1.613***		-1.610***	
	(0.0464)		(0.0466)	
Competitor	-0.00957***	0.00234***	-0.00957***	0.00227***
	(0.00152)	(0.000575)	(0.00152)	(0.000574)
$ln(Seller\ Reputation)$	0.0236**	0.00972**	0.0239**	0.0103*
	(0.0115)	(0.00401)	(0.0115)	(0.0039)
$Seller's\ Experience$	5.80e-05	-3.42e-05*	5.74e-05	-3.30e-05*
	(0.000105)	(1.98e-05)	(0.000105)	(1.98e-05)
Seller is Dealer	-0.322***		-0.319***	
	(0.0406)		(0.0405)	
$Warranty$	-0.142***	0.0968***	-0.142***	0.0938***
	(0.0447)	(0.0181)	(0.0446)	(0.0180)
$AMileage$	0.000913	-0.00101***	0.000911	-0.000992***
	(0.00919)	(0.000245)	(0.00921)	(0.000245)
$Duration = 3Days$	-0.0331	-0.0839***	-0.0330	-0.0815***
	(0.0939)	(0.0296)	(0.0938)	(0.0295)
$Duration = 5Days$	0.0743	-0.0466**	0.0741	-0.0436
	(0.0656)	(0.0222)	(0.0656)	(0.0221)
$Duration = 7Days$	-0.00750	-0.0158	-0.0754	-0.0151
	(0.0519)	(0.0183)	(0.0519)	(0.0182)
Constant	0.803**	0.511***	0.800*	0.499***
	(0.4070)	(0.0310)	(0.4060)	(0.0311)
Fuel Type	Y	N	Y	N
Car Model	Y	N	Y	N
Car Body Type	Y	N	Y	N
Vehicle Condition	Y	N	Y	N
ath Rho		0.482***		0.500***
		(0.0751)		(0.0755)
$ln(Sigma)$		-1.334***		-1.334***
		(0.0211)		(0.0213)
Observations	10,893	10,893	10,893	10,893

Table 9: Average Shill Scores for the Bidders

Shill Score	Criteria: 5%	Criteria: 30 minutes	Criteria: 5 minutes
Before threshold	0.443	0.436	0.434
	(58,240)	(65,781)	(67,834)
After threshold	0.359	0.338	0.322
	(16,478)	(8,006)	(5,833)

The number in parentheses is the number of bidders.

Table 10: Proportions of Shill Bidders before and after Threshold

Shill Score	Criteria: 5%	Criteria: 30 minutes	Criteria: 5 minutes
Before threshold	3.25%	3.00%	2.96%
	(1,895)	(1,974)	(2,006)
After threshold	1.12%	1.17%	1.14%
	(185)	(94)	(67)

The number in parentheses is the number of bidders.

Figure 1: The Distribution of Bidders' Shill Scores

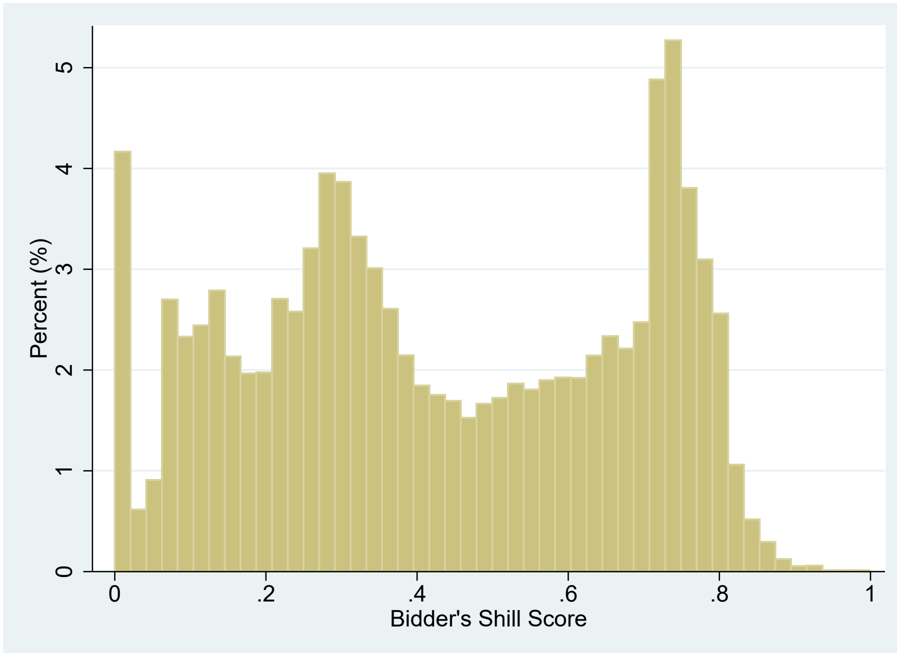

Figure 2: The Distribution of Average Shill Scores

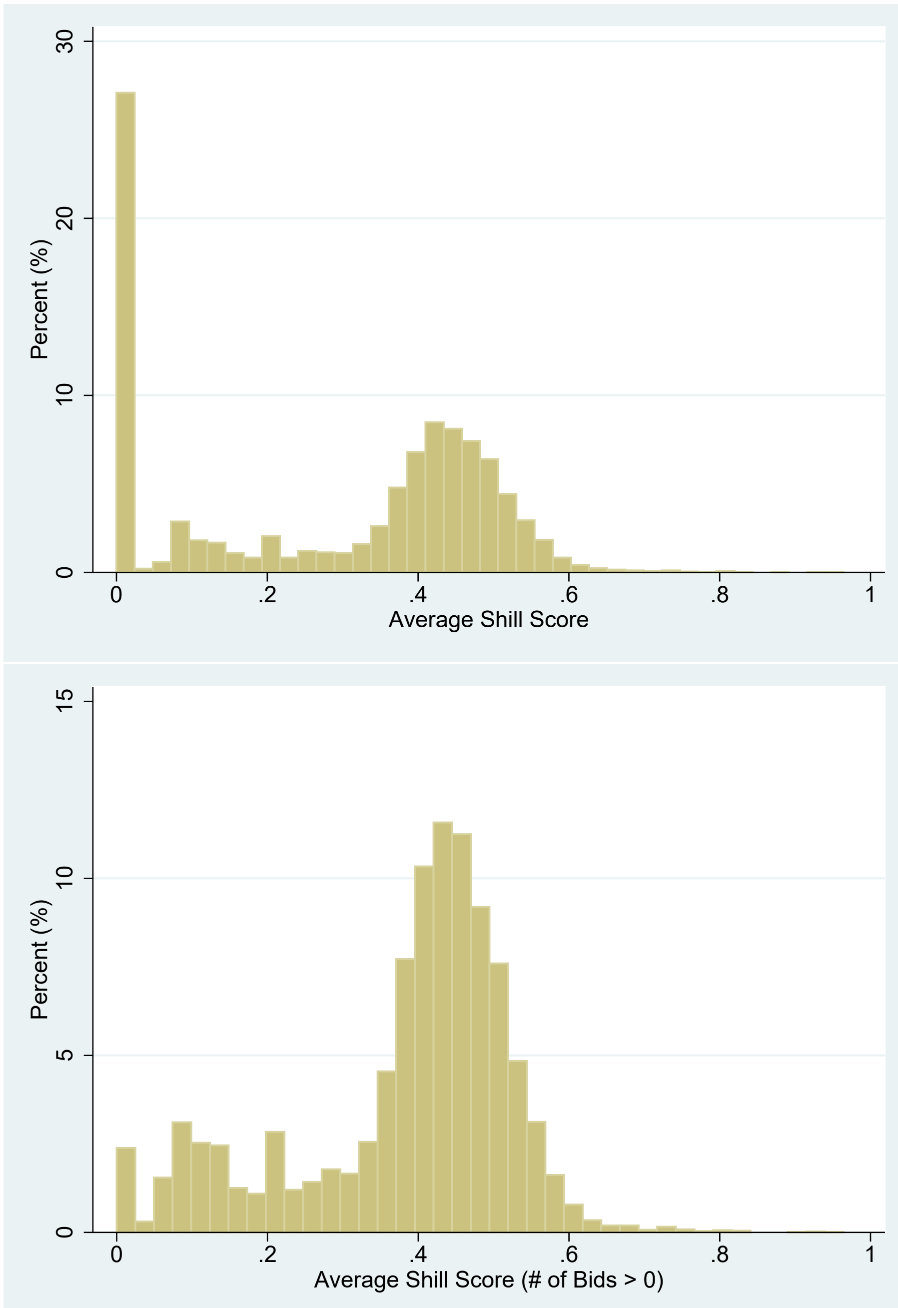

Figure 3: Comparison of Trade Probability, Transaction Price, and Expected Revenue