Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

How Reduced Search Costs and the Distribution of Bidder Participation Affect Auction Prices

Eric Overby, Karthik Kannan

MANAGEMENT SCIENCE
Vol. 61, No. 6, June 2015, pp. 1398–1420
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2014.2043

How Reduced Search Costs and the Distribution of Bidder Participation Affect Auction Prices

Eric Overby
Scheller College of Business, Georgia Institute of Technology, Atlanta, Georgia 30308, eric.overby@scheller.gatech.edu

Karthik Kannan
Krannert School of Management, Purdue University, West Lafayette, Indiana 47904, kkarthik@purdue.edu

Electronic commerce allows bidders to find and participate in auctions regardless of location. This reduction in bidders' search costs has important effects on bidders' participation patterns and sellers' revenue. The "demand expansion" effect occurs when reduced search costs allow bidders to participate in more auctions. The "demand distribution" effect occurs when reduced search costs allow bidders to distribute themselves more evenly across auctions. We focus on the latter effect by modeling when a more even distribution of bidder participation across auctions increases seller revenue. We apply our analytical insights to 65,718 sequential auctions (comprising over 10 million vehicles) in the wholesale used vehicle market. We show that reduced search costs can increase seller revenue by smoothing the distribution of bidder participation across auctions, even if the aggregate amount of bidder participation remains constant. This contributes new results to the auction theory literature and generates novel insights for sellers seeking increased revenue.

Keywords: electronic commerce; search costs; multiple-object auctions; distribution of bidder participation
History: Received December 22, 2011; accepted July 14, 2014, by Lorin Hitt, information systems. Published online in *Articles in Advance* February 23, 2015.

1. Introduction

Participation in an auction has traditionally required bidders to be collocated at an auction hall or similar facility. The prices that sellers could fetch for their objects were limited by the number of bidders they could attract to these facilities. Advances in information technology have allowed bidders to participate in auctions electronically. For example, eBay is well known for its Internet auctions, and many industries in which transactions are conducted via auction have implemented electronic auction systems (e.g., Koppius et al. 2004). These systems allow bidders to identify and participate in auctions with the click of a mouse, thereby lowering their search costs. We examine how the reduction in search costs due to these information systems affects prices. We consider two effects, which we label "demand expansion" and "demand distribution." First, reduced search costs might enable bidders to participate in more auctions, which could lead to higher prices due to increased bidding competition. Although we believe this demand expansion effect to be interesting, it is also rather intuitive, as it essentially reduces to the argument that increased demand leads to higher prices (e.g., Lee 1998). Thus, rather than focusing on this effect, we focus on a more nuanced and (we believe) surprising effect that we label the demand distribution effect. We argue that lower search costs may lead to a more even distribution of bidder participation across auctions, because bidders with low search costs can easily monitor multiple auctions and shift from auctions with high bidding competition to those with low competition. We show analytically that a more even distribution often yields higher prices, even if aggregate demand remains constant, i.e., if there is no demand expansion. We show that this holds under fairly general conditions, but that it depends on the variability of bidder valuations for the objects being auctioned, the degree to which sellers can (and do) predict this a priori, and the average number of bidders per auction.

The effect of the distribution of bidder participation across auctions on prices has received little research attention. Understanding this effect is important because advances in information technology continue to improve bidders' ability to participate in auctions regardless of location, leading to interesting changes in bidder participation patterns. Our analysis of how these changes affect auction prices contributes new results to the auction theory literature, and it generates novel insights for sellers seeking increased revenue. For example, a standard strategy for a seller seeking increased revenue is to increase bidder participation in her auctions. We show that even if a seller is unable to increase bidder participation, she can still increase her revenue by implementing strategies to engender a more even distribution of bidder participation across auctions (e.g., by spreading auctions across time or space, by using promotions to shape bidder participation, etc.).

We apply the insights of our model to the empirical context of the wholesale used vehicle market. We use a unique data set of over 65,000 sequential auctions (comprising over 10 million used vehicles) conducted from 2005 to 2010 in this market to study how the distribution of bidder participation across these auctions influences the high bids. The sequential auctions in this market are held at physical facilities, and bidders are used car dealers purchasing inventory for their dealerships. An important feature of this empirical context is that the physical auctions are "simulcast'' on the Internet, which allows bidders to participate either physically or electronically. We show that physical bidders tend to bid early in the sequence, whereas electronic bidders tend to bid later in the sequence, which we attribute to differences in the search costs faced by bidders when using the two channels. The later bidding of electronic bidders smoothes the distribution of participation across the auctions in the sequence. We show that a more even distribution of bidder participation across auctions is positively associated with the high bids received in the auctions. Consistent with the analytical model, this relationship is moderated by the average number of bidders who participate in the auctions and whether the seller sequences her vehicles based on the variability of bidder valuations.

2. Literature Review

Our research contributes to two streams of literature: (a) the literature on the Internet, search costs, and prices; and (b) the literature on multiple-object auctions.

2.1. The Internet, Search Costs, and Prices

Studies in information systems, marketing, and economics have examined how the Internet reduces search costs and the effect this has on prices. Smith et al. (1999) defined search costs as being comprised of external and internal search costs. External search costs include the monetary and opportunity costs associated with gathering information, and internal search costs include the cognitive effort expended while gathering and processing the information.

Several analytical models have been proposed to link search costs to prices. Baye et al. (2006) grouped these models into two categories. First, *search theoretic models* assume that each buyer pays a cost to obtain a price quote from each seller (e.g., Stahl 1989). "Fixed sample search" models assume that buyers commit to searching a fixed number of sellers, and "sequential search" models assume that buyers continue searching until the cost of an additional search exceeds its benefit. Second, *information clearinghouse models* assume that a list of prices is aggregated in a repository such as a newspaper or Internet price comparison site (e.g., Varian 1980). Motivated by these models, several empirical studies have examined how search costs affect prices, with a recent focus on whether the Internet has affected prices by reducing search costs (Brown and Goolsbee 2002, Brynjolfsson and Smith 2000, Clemons et al. 2002). Whether reduced buyer search costs lead to lower or higher prices is equivocal. Empirical findings are mixed, and analytical predictions depend on model assumptions (see Baye et al. 2006).

Much of this literature assumes that sellers set prices. For example, Baye et al. (2006) provided an impressive review of the literature in this stream; all of the models they reviewed involve sellers who set prices. By contrast, we consider the case in which prices are determined by auction. Whether prices are posted by the seller or determined by auction has important implications for how search affects prices. The reason is simple yet fundamental. In a posted price setting, the outcome of a buyer's search process (i.e., his choice of a seller) does not affect the seller's price for that transaction (although the seller might adjust her posted price for future transactions). This is not true in an auction setting, where a buyer's choice of a seller may affect the transaction price through changes in bidding competition.

2.2. Multiple-Object Auctions

Multiple-object auctions refer to situations in which a seller(s) auctions more than one object that may or may not be identical. There are many possibilities for conducting multiple-object auctions: the objects may be auctioned as a bundle in a single auction, each object may be auctioned individually in multiple auctions, or a combination of the two may be used. If multiple auctions are conducted, then they may be conducted sequentially, simultaneously, or in an overlapping manner. Scholars in information systems and other disciplines have made several contributions to our understanding of multiple-object auctions. This includes recent research on how different policies for revealing information about prior auctions influence bidder learning and auction outcomes (e.g., Arora et al. 2007, Greenwald et al. 2010, Kannan 2012) and how bidders' willingness to pay evolves in sequential auctions for identical objects (Goes et al. 2010). Other research has focused on overlapping auctions, including how factors such as the degree of overlap and the information revealed about prior auctions affect price (Bapna et al. 2009), as well as simultaneous auctions, including how bidders with unit demand choose the auction in which to participate (Bapna et al. 2010). Other research has examined how sellers use multiple auctions as a means of searching for high-valuation buyers (Genesove 1995, Kuruzovich et al. 2010). Our analysis contributes to this stream by showing how the distribution of bidder participation across auctions in a multiple-object setting affects prices.

3. Theory

We conceptualize the process of a bidder observing or participating in multiple auctions as him searching for the auction(s) that provides him the most surplus. Bidders must balance the benefits of continuing to search with the costs of doing so (as in a classical sequential search model). Traditionally, bidders have participated in auctions physically, although technological advances have provided them the opportunity to participate electronically. The search cost that a bidder incurs to observe (and potentially participate in) each auction varies based on whether he is using a physical or electronic channel. As is commonly assumed in the literature on search costs in physical and electronic environments (Baye et al. 2006), we assume that search costs are lower in the electronic channel. This allows bidders using the electronic channel to observe/participate in more auctions than can bidders using the physical channel. (For example, in a sequential auction setting, electronic bidders can arrive earlier in the sequence and/or stay later.) This allows them to be more selective and increases the likelihood that they will participate in auctions with low competition instead of auctions with high competition, assuming that the auctions are for substitutable objects. We refer to these shifts from high competition to low competition auctions as Robin Hood operations (Arnold 1987).[1] All bidders seek to conduct Robin Hood operations, ceteris paribus, because that will expose them to less bidding competition and provide them with greater surplus. However, physical bidders have less ability to conduct Robin Hood operations than do electronic bidders (on average) because the former cannot observe as many auctions. In some cases, a physical bidder will participate in an auction even if he realizes that bidding competition is high, because participation still provides him with positive surplus and it is too costly to wait for a later auction. As a result, having a mix of bidders with high and low search costs should result in a more even distribution of bidder participation across auctions than having only bidders with high search costs. We explore this in our empirical context in §4.2. Under fairly general conditions (which we detail below), a more even distribution of bidder participation across auctions will lead to higher expected revenue for the seller.

Before proceeding, we note the potential for confusion regarding the term "distribution." When we say "distribution of bidder participation," we are referring to the manner in which bidder participation is spread across auctions. We will also use the term "distribution" to refer to probability distributions.

3.1. Model Setup

Consider a multiple-object auction scenario i in which J auctions are conducted, referenced by $j = 1, 2, \ldots, J$. Suppose there are M objects that may be auctioned in scenario i, referenced by $m = 1, 2, \ldots, M$. Let p_{jm} represent the probability of object m being offered in auction j. This probability allows us to model the seller's ability to allocate certain objects to certain auctions. In any auction j, a single object m is auctioned (equivalently, $\sum_{m=1}^{M} p_{jm} = 1\ \forall\, j$). Let $F_m(\cdot)$ represent the probability distribution from which bidder valuations for object m are independently drawn. As is commonly assumed in the literature, $F_m(\cdot)$ is a "regular" distribution (Myerson 1981).[2] Let x_{ij} represent the number of bidders who participate in auction j in scenario i. The same bidder can participate in more than one auction and can purchase more than one object. Collect the x_{ij} values in a vector labeled X_i, i.e., $X_i = (x_{i1}, x_{i2}, \ldots, x_{iJ})$. Let the aggregate amount of bidder participation in scenario i be $X_i^{\text{total}} = \sum_{j=1}^{J} x_{ij}$.

We simplify our model by ignoring the learning of bidder valuations across auctions. This simplification is justified as follows. First, theoretical analysis of multiple-object auctions has often focused on the sequential auction of two identical objects in which bidders learn each other's valuations for the objects (e.g., Bernhardt and Scoones 1994, Kannan 2010). By contrast, we allow there to be more than two auctions. Solving for the equilibrium in a general multiperiod setting is hard because it involves modeling how each bidder solves his own multidimensional dynamic optimization problem as a best response to how all other bidders solve their problems (Brendstrup and Paarsch 2006). Given this difficulty, it is unreasonable

[1] The "Robin Hood operation" term was coined as a synonym for progressive transfers in income inequality economics in which income is shifted from a rich person to a poor person, with the magnitude of the shift being less than the original difference between the persons' incomes (Arnold 1987). We adapt it for our purposes to describe instances in which bidders shift from high competition auctions (i.e., "rich" auctions) to low competition auctions (i.e., "poor" auctions). For example, assume that eight bidders participate in auction A and four bidders participate in auction B. A bidder who shifts from auction A to auction B conducts a Robin Hood operation.

[2] Because the distribution is regular, its hazard rate is monotonically increasing (Myerson 1981). Many common distributions have an increasing hazard rate, including normal, uniform, logistic, extreme value, and χ^2.

to expect bidders to compute and play their optimal strategies. Furthermore, the equilibrium so computed will likely not be stable because, for it to be stable, *every* bidder has to compute and play his equilibrium strategy. This creates doubt as to whether a model that incorporates bidder learning would accurately represent actual bidder behavior. Second, and as a consequence of the previous point, simplifications in which bidder learning is not modeled have precedence in the literature (e.g., see Edelman et al. 2007). Third, our analysis involves *nonidentical* objects; as such, the information a bidder learns about other bidders' valuations for one object may be irrelevant to his bidding strategy for a different object. Multiple-object auctions in which the objects are nonidentical are common in practice, as many traditional auctions for collectibles, antiques, and livestock are conducted this way. Given the above, we assume that valuations for each object are drawn independently of the auction in which it is offered. Hence, the expected price for object m in auction j in scenario i is equal to the valuation of the second-highest bidder in that auction. If object m attracts x_{ij} bidders, then its expected price, $Price_m(x_{ij})$, is the expectation of the $(x_{ij}-1)$st order statistic from the set of the x_{ij} bidders' valuations. As shown in Lemma A1 in the appendix, $Price_m(x_{ij})$ is concave in x_{ij} given that $F_m(\cdot)$ is a regular distribution. Let $TotalPrice(X_i)$ be the total expected revenue for the seller in scenario i. Because $TotalPrice(X_i)$ is separable in j, $TotalPrice(X_i)=\sum_{m=1}^{M}\sum_{j=1}^{J}(p_{jm}*Price_m(x_{ij}))$.

Consider two scenarios, 1 and 2, in which a seller conducts the same J auctions for the same M objects. X_1 and X_2 represent the vector of x_{ij} values for the respective scenarios. Assume that the aggregate amount of bidder participation in the J auctions is the same in both scenarios, i.e., $X_1^{\text{total}}=X_2^{\text{total}}$. This means that there is no demand expansion from one scenario to the other. The two scenarios differ only in how bidder participation is distributed across the J auctions, thereby allowing us to isolate the effects of demand distribution, which is the focus of our analysis. Let scenario 1 involve only physical bidders and result in the distribution of bidder participation represented by X_1. Let scenario 2 mirror scenario 1 except that some bidders switch to the electronic channel. The lower search costs for electronic bidders in scenario 2 mean that they are more likely to conduct Robin Hood operations in which they shift from a high competition auction to a low competition auction, as discussed above. Thus, the distribution of bidder participation represented by X_2 results from a series of Robin Hood operations on X_1. As shown by Arnold (1987, Chap. 2, pp. 11–13), these Robin Hood operations mean that X_1 majorizes X_2 and that the variance of the elements in X_1 exceeds that of X_2.[3] Therefore, the distribution of bidder participation in scenario 2 is more even than that in scenario 1. Because we want to examine how the evenness of the distribution of bidder participation across auctions affects prices, we compare $TotalPrice(X_1)$ to $TotalPrice(X_2)$.

3.2. Model Analysis

We use two models to analyze when a more even distribution of bidder participation yields higher expected revenue. The first is for the general case involving J auctions. The second is for a simple case involving two auctions, which we use to explore aspects of the problem that are intractable in the general case.

3.2.1. Model 1: General Case Involving J Auctions. Following Engelbrecht-Wiggans's (1994) notion of stochastic equivalence, we begin with the case in which $p_{jm}=p_m$, where p_m denotes the probability of object m appearing in any given auction j. (We relax this restriction later.) The following theorem holds.

Theorem 1. *With stochastically equivalent auctions, total expected seller revenue is higher (or the same) in scenario 2 (in which bidder participation is more evenly distributed across auctions) than in scenario 1 (in which bidder participation is less evenly distributed across auctions).*

Proof. Because $p_{jm}=p_m$, $TotalPrice(X_i)=\sum_{m=1}^{M}p_m\cdot\sum_{j=1}^{J}Price_m(x_{ij})$. Because $Price_m(x_{ij})$ is a concave function (see Lemma A1) and X_1 majorizes X_2, it follows from Arnold (1987, Theorem 2.9, p. 24) that $-\sum_{j=1}^{J}Price_m(x_{1j})\geq-\sum_{j=1}^{J}Price_m(x_{2j})$. Thus, $p_m\sum_{j=1}^{J}Price_m(x_{1j})\leq p_m\sum_{j=1}^{J}Price_m(x_{2j})$. Summing both sides over all m, we have $\sum_{m=1}^{M}p_m\cdot\sum_{j=1}^{J}Price_m(x_{1j})\leq\sum_{m=1}^{M}p_m\sum_{j=1}^{J}Price_m(x_{2j})$, which means $TotalPrice(X_1)\leq TotalPrice(X_2)$. QED[4]

[3] For vector $X_i=(x_{i1},x_{i2},\ldots,x_{iJ})$, let $X_i^{\downarrow}$ contain the same elements as X_i sorted in descending order. X_1 "majorizes" X_2 if $\sum_{j=1}^{k}x_{1j}^{\downarrow}\geq\sum_{j=1}^{k}x_{2j}^{\downarrow}$ for $k=1,2,\ldots,J$ and $\sum_{j=1}^{J}x_{1j}=\sum_{j=1}^{J}x_{2j}$. For example, let $J=3$, let X_1 be (6, 4, 2), and let X_2 be (5, 4, 3). Note that X_1 is more unevenly distributed than X_2 and that X_1 majorizes X_2. A Robin Hood operation on $X_1=(6, 4, 2)$ in which 1 is subtracted from the 1st element and added to the 3rd element results in $X_2=(5, 4, 3)$. This illustrates why Robin Hood operations result in more evenness of the elements in a vector.

[4] Suppose the auction's starting price means that only the bidders whose valuations lie to the right of a truncation point of the original valuation distribution are observed to bid. We prove here that the truncated distribution retains the increasing hazard rate property of the original distribution. Let $f(x)$ and $F(x)$, respectively, be the probability density function (pdf) and cumulative distribution function (cdf) of the original valuation distribution with support $[\underline{v},\bar{v}]$. Let $g(x)$ and $G(x)$, respectively, be the pdf and cdf of the truncated distribution with support $[a,\bar{v}]$, where a corresponds to the truncation point. By definition, $g(x)=f(x)/(1-F(a))$ and $G(x)=(F(x)-F(a))/(1-F(a))$. Because the original valuation distribution is regular, it has the increasing hazard rate property

To expound on the above result, note that p_{jm} determines how the seller assigns the M objects to the J auctions. The seller might strategically set p_{jm} using knowledge of the distribution of bidder valuations for the objects being auctioned (i.e., the $F_m(\cdot)$ distributions). For example, if certain auctions tend to attract a high number of bidders (which might be the case, say, for the initial auctions in a sequential auction setting), then the seller might be more likely to offer objects with a wide range of bidder valuations in those auctions, because having a high number of bidders will increase price more when bidder valuations are widely dispersed than when they are narrowly dispersed. However, this assumes that the seller (a) knows the distribution of bidder valuations for her objects and (b) uses that knowledge to allocate objects to auctions.[5] Often, this will not be the case, such that modeling $p_{jm} = p_m$ is reasonable and Theorem 1 will apply. We discuss this further in §§3.2.2 and 3.2.3.

Next, we drop the $p_{jm} = p_m$ restriction, thus allowing the seller to allocate objects to auctions based on knowledge of the distribution of bidder valuations, although we add restrictions on the $F_m(\cdot)$ distributions. We allow the support of the $F_m(\cdot)$ distributions (denoted as $[\underline{v}_m, \bar{v}_m]$) to vary across objects m, but we impose the restriction that the range (denoted by $r_m = \bar{v}_m - \underline{v}_m$) and shape of the $F_m(\cdot)$ distributions be the same. (We consider relaxations of these restrictions later in model 2.) The following holds.

Proposition 1. *If bidder valuations for each auctioned object are drawn from distributions with the same range and shape, but potentially with different supports, then total expected seller revenue is higher in scenario* 2 (*in which bidder participation is more evenly distributed across auctions*) *than in scenario* 1 (*in which bidder participation is less evenly distributed across auctions*).

Proof. The proof is similar to that for Theorem 1. Because $\bar{v}_m - \underline{v}_m = r\ \forall\, m$, define $F_0(\cdot)$ as $F_m(\cdot)$ shifted to the left such that $[\underline{v}_m, \bar{v}_m]$ becomes $[0,\ r]$. Define $\varphi(x_{ij})$ as the expectation of the second-highest order statistic when the bidder valuations are drawn from $F_0(\cdot)$ in auction j in scenario i. Thus, $Price_m(x_{ij}) = \underline{v}_m + \varphi(x_{ij})$. Therefore, $TotalPrice(X_i) = \sum_{m=1}^{M}\sum_{j=1}^{J}[p_{jm} * (\underline{v}_m + \varphi(x_{ij}))] = \sum_{m=1}^{M}\underline{v}_m\sum_{j=1}^{J}p_{jm} + \sum_{j=1}^{J}\varphi(x_{ij})$ (recall that $\sum_{m=1}^{M}p_{jm} = 1\ \forall\, j$). Because $\varphi(x_{ij})$ is concave per Lemma A1, we can complete the proof as in Theorem 1. QED

3.2.2. Model 2: Simple Case Involving Two Auctions. Consider two nonidentical objects (A and B) that are offered, individually, in two auctions (1 and 2). Either object can be offered in either auction. Each bidder for object A (B) draws a valuation from a uniform distribution with support $[\underline{v}_a, \bar{v}_a]$ ($[\underline{v}_b, \bar{v}_b]$).[6] Let $r_a = \bar{v}_a - \underline{v}_a$ with r_b analogous. Without loss of generality, set $r_a \geq r_b$. i.e., the range (and hence the shape) of bidder valuations can differ. Consider two scenarios. The aggregate amount of bidder participation is the same in both scenarios (let it be equal to $2\dot{x}$), but the distribution of bidder participation differs; i.e., there is no demand expansion between the scenarios, only differences in demand distribution, as above. In scenario 1, we assume $\bar{x}_1 = \dot{x} + \alpha$ bidders in auction 1 and $\underline{x}_1 = \dot{x} - \alpha$ in auction 2. In scenario 2, we assume $\bar{x}_2 = \dot{x} + \beta$ bidders in auction 1 and $\underline{x}_2 = \dot{x} - \beta$ bidders in auction 2. Note that $\dot{x}$ is the average number of bidders in the auctions, and $\alpha > \beta \geq 0$ represents the degree to which the distribution of bidder participation across the auctions is uneven. The condition $\alpha > \beta$ means that the distribution of bidder participation is more even in scenario 2 than in scenario 1.[7] $\beta = 0$ represents the special case in which the distribution of bidder participation in scenario 2 is perfectly even. If the seller assigns the object with the widest range of bidder valuations (object A) to the auction that attracts more bidders (auction 1), then this could increase her overall revenue, as noted above. To do this, however, she must know the distribution of bidder valuations for the objects a priori, such that she can identify object A as having the wider range of valuations. Let p_{1A} (which mirrors p_{jm} from model 1) represent the probability with which she does this.

Proposition 2. *Total expected seller revenue is higher in scenario* 2 (*in which bidder participation is more evenly distributed across auctions*) *than in scenario* 1 (*in which bidder participation is less evenly distributed across auctions*) *when either of the following sufficient* (*but not necessary*) *conditions holds*:

(a) $p_{1A} \leq 0.5$. *i.e., the seller cannot* (*or does not*) *identify a priori the object with the widest range of bidder valuations any better than chance.*

$f(x)/(1 - F(x)) > 0$. Using the above relations, it can be shown that $g(x)/(1 - G(x)) > 0$, i.e., that the truncated distribution retains the increasing hazard rate property. Therefore, the concavity result from Lemma A1 holds for the truncated distribution, which means that Theorem 1 and subsequent proofs also hold for the truncated distributions as long as the truncation points are consistent across auctions.

[5] This would require a nontrivial level of seller rationality, especially because the distributions may change with evolving market conditions and buyer preferences.

[6] The assumption of a uniform distribution simplifies the analysis and satisfies the requirement that the distribution of bidder valuations follows a regular distribution. Similar results apply to other regular distributions, although the region defined in Proposition 2 is specific to a uniform distribution.

[7] The variables $\bar{x}_1$, $\underline{x}_1$, $\bar{x}_2$, and $\underline{x}_2$ are positive integers to ensure that someone is bidding in each auction. When coupled with $\alpha > \beta \geq 0$, this implies that $\dot{x} \geq 2$.

(b) $r_a = r_b$; *i.e., the range of bidder valuations is the same for both objects.*

If neither condition holds, then there is a region in which the more even distribution yields higher total expected revenue, with the region defined by the following expression:

$$(\alpha+\beta)(\dot{x}+1)(r_a+r_b) > (\alpha\beta+(\dot{x}+1)^2)(2p_{1A}-1)(r_a-r_b). \quad (1)$$

Proof. Consider the case when neither conditions (a) nor (b) hold. The expected price is the expectation of the second-highest order statistic from the set of bidder valuations. From the properties of a uniform distribution bounded by $[\underline{v}, \bar{v}]$, the expected value of this order statistic given x bidders is $\underline{v}+((x-1)/(x+1))(\bar{v}-\underline{v})$. In scenario 1, the expected total revenue from the two auctions is $p_{1A}(\underline{v}_a+((\bar{x}_1-1)/(\bar{x}_1+1))(r_a)+\underline{v}_b+((\underline{x}_1-1)/(\underline{x}_1+1))(r_b))+(1-p_{1A})(\underline{v}_a+((\underline{x}_1-1)/(\underline{x}_1+1))(r_a)+\underline{v}_b+((\bar{x}_1-1)/(\bar{x}_1+1))(r_b))$. In scenario 2, the expected total revenue is analogous, with $\bar{x}_2$ and $\underline{x}_2$ replacing $\bar{x}_1$ and $\underline{x}_1$. The proof is accomplished by determining when the total expected revenue for scenario 2 is greater than or equal to the total expected revenue for scenario 1. After algebraic manipulation, this reduces to (1). Note that the left-hand side of (1) is always positive. Result (a) can be verified by checking that, for $p_{1A} \le 0.5$, the right-hand side of (1) is always less than or equal to 0. Result (b) can be verified by noting that the right-hand side is equal to 0 when $r_a = r_b$. QED

Proposition 2 yields several insights about our overall model. First, as r_a gets closer to r_b, the right-hand side of (1) gets smaller, making (1) more likely to hold. This illustrates that similarity among the range and shape of the bidder valuation distributions (rather than equivalence) may be sufficient for a more even distribution of bidder participation to yield higher total revenue. Second, even if $p_{1A}=1$ (i.e., the seller perfectly identifies the object with the widest range of bidder valuations), (1) will often still hold, depending on the values of the other parameters. Third, a more even distribution is more likely to yield higher total expected revenue when the average number of bidders ($\dot{x}$ in the model) is small. Although this holds for any value of α and β, it is easy to see when $\alpha=1$ and $\beta=0$. In that case, (1) reduces to $r_a+r_b \ge (\dot{x}+1)(2p_{1A}-1)(r_a-r_b)$. As x grows, it becomes less likely that (1) will hold.[8]

[8] Another way to see this is to create a gap function by subtracting the right-hand side of (1) from the left-hand side, differentiating with respect to $\dot{x}$, and setting equal to 0. This yields $(\alpha+\beta)(r_a+r_b)-(2\dot{x}+2)(2p_{1A}-1)(r_a-r_b)=0$. We know that $r_a-r_b>0$ by construction. We also know that the proof holds for $p_{1A}\le 0.5$, so we need only concern ourselves with the case where $p_{1A}>0.5$, such that $2p_{1A}-1>0$. In that case, the gap function is always decreasing in $\dot{x}$, which means that (1) is less likely to hold at higher values of $\dot{x}$.

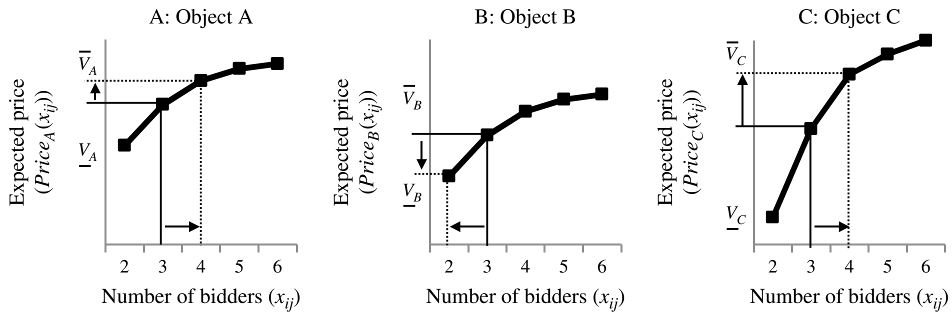
Proposition 2 also illustrates that the object-to-auction allocation method that is relevant for our theory is that the seller allocates objects based on the distribution of bidder valuations; we account for this via p_{1A} (and p_{jm} in the general case). Other allocation methods, such as allocation based on object quality or supply, will not affect our result about how a more even distribution of bidder participation can increase total expected revenue—unless the other methods are surrogates for allocation based on the bidder valuation distributions; i.e., we do not require that the seller allocate objects randomly.

3.2.3. Summary and Intuition for the Results. As illustrated in Figure 1, because $Price_m(x_{ij})$ is concave, adding a bidder to an auction adds less to the price than removing a bidder subtracts from the price. Thus, if the range and shape of the $F_m(\cdot)$ distribution are the same for the objects being auctioned, then distributing bidders evenly across auctions will yield higher expected revenue than distributing them unevenly; see Proposition 1. This is illustrated in panels A and B of Figure 1; notice that a bidder distribution of (3, 3) will yield higher revenue than a distribution of (4, 2) or (2, 4). This result will also hold if the ranges and shapes of the $F_m(\cdot)$ distributions are similar (as opposed to equivalent) for the objects being auctioned, with the necessary "similarity" characterized in Proposition 2 for the two-auction case. If the variability of the $F_m(\cdot)$ distributions (as reflected in their range and shape) differs substantially across objects, then shifting a bidder from an object with a low variability of bidder valuations to an object with a high variability may yield a net increase in total expected revenue, even if this shift results in a more uneven distribution of bidder participation (as illustrated in panels B and C of Figure 1). However, this strategy can only be implemented if the seller can identify (a priori and with some probability greater than chance) the object(s) with the highest variability of bidder valuations. If the seller cannot (or does not) do this, then her expected revenue will be higher with a more even distribution of bidder participation; see Theorem 1 and Proposition 2. Because $Price_m(x_{ij})$ is increasing and concave in x_{ij}, adding or removing a bidder has a larger impact on $Price_m(x_{ij})$ when x_{ij} is small than when x_{ij} is large. This can be seen by the steeper slope of the $Price_m(x_{ij})$ function at smaller values of x_{ij} (see Figure 1). This means that a more even distribution of bidders is more desirable for the seller if the average number of bidders is low; see Proposition 2.

4. Empirical Application

The insights from our theoretical analysis are applicable to any context in which sellers use multiple

Figure 1 **Graphical Intuition for the Analytical Results**



Notes. Each chart shows the $Price_m(x_{ij})$ function for three different objects m: A, B, and C. The range of the bidder valuations for each object, $\bar{v}_m - \underline{v}_m$, is depicted on the y axis. The solid lines depict the expected price when $x_{ij} = 3$. The arrows and dashed lines depict the change in expected price after adding or removing a bidder.

auctions to sell objects, including markets for livestock, agricultural goods, industrial equipment, collectibles, and art (e.g., Ashenfelter 1989, Kazumori and McMillan 2005). Here, we apply our theoretical analysis and illustrate its implications in the context of the wholesale used vehicle market.

4.1. Market Background

The wholesale used vehicle market is a business-to-business market in which buyers and sellers trade used vehicles. The buyers are used car dealers purchasing inventory for their retail lots. Sellers are either (a) other dealers, who sell vehicles that they do not wish to sell on their retail lots, or (b) institutional sellers such as rental car companies (who sell vehicles retired from their rental fleets). In 2011, 7.9 million used vehicles were sold in the U.S. wholesale market for a total transaction value of approximately \$73 billion (NAAA 2012).

The market in the United States has traditionally functioned as follows. There are multiple automotive auction companies that broker transactions between buyers and sellers. Sellers transport vehicles to facilities operated by these companies. Vehicles are auctioned sequentially in a "sales event." The sequence, including the number of vehicles to be auctioned, is published before the sales event in the "run list." Each vehicle in a sales event is driven, one at a time, into a warehouse-type building, where an auctioneer solicits bids on the vehicle. The auction format is an open outcry English auction. The bidding for each vehicle takes approximately 30–45 seconds, after which the next vehicle is auctioned. Because each used vehicle is different, each sales event represents a sequential auction for nonidentical objects. An entire sales event can last several hours. The position in the sequence in which each vehicle is auctioned is referred to as the "run number." Generally speaking, there are two types of sales events that reflect the two types of sellers in the market: sales events for dealer sellers and sales events for institutional sellers. A dealer sales event consists of vehicles offered by more than one dealer seller (typically), whereas an institutional sales event consists of vehicles offered by a single institutional seller. For example, a dealer sales event might consist of 150 vehicles being sold by 10 different dealers (each offering 15 vehicles), whereas an institutional sales event might consist of 150 vehicles being sold by Avis Rent A Car. No two vehicles in a sales event are the same; even vehicles of the same year/model differ due to color, option packages, trim level, mileage, and wear and tear. These differences mean that even if a bidder learns another bidder's valuation for a vehicle, that information may be irrelevant to his bidding strategy for another vehicle, even if it is of the same year/model. Also, because buyers in this market are acquiring inventory for resale to retail customers, they are often willing to substitute one vehicle for another. This is consistent with the theoretical model (see §3).

Bidders can bid on each vehicle in one of two ways. They can travel to the physical facilities and place bids in person, or they can place bids electronically via a web-based application that "simulcasts" live audio and video of sales events as they occur at the physical facility. Thus, each auctioned vehicle may be bid upon by bidders using the physical channel and bidders using the electronic "webcast" channel.

As in §3, we conceptualize a bidder's continued observation/participation in the auctions in each sales event as a search for the auction(s) that provides him the most surplus. As the auctions progress, each bidder can observe the bidding competition in each auction, assess whether it is high or low, and decide whether to participate in that auction or wait for a later auction. Although waiting may allow the bidder to participate in an auction with less competition, he incurs the cost of waiting, i.e., the search cost. This cost is higher for physical bidders than for electronic bidders for the following reasons. First, physical bidders must pay the opportunity costs (which are a component of search costs; see §2) of the time spent away from their dealerships, which accumulate with each successive auction. These costs are lower for electronic bidders because they can bid on vehicles while performing other tasks at their dealerships. Second, it is more physically costly to continue to participate in a sales event when at the physical facility than when using the electronic channel. This is because the physical facility requires bidders to stand while bidding, the air quality is relatively poor given exhaust fumes in the facility, the facility is loud, etc.

4.1.1. Raw Data and Variables. We obtained data from an automotive auction company. The data consist of 10,351,857 vehicles auctioned in 65,718 institutional sales events in the United States from January 11, 2005, to March 31, 2010. Complete data for the period from June 1, 2007, to December 31, 2008, were not available because the bid logs could not be retrieved.[9] We focused on institutional sales events for the following reasons. First, because institutional sales events have only one seller, we can include seller fixed effects in our regressions to control for seller characteristics such as size, reputation, geographic presence, etc. Second, because an institutional seller has control over all vehicles in her sales event, we can (and do) estimate whether the seller sequences her vehicles based on the variability of bidder valuations, which is one of the variables identified in the analytical model. We also examine whether the seller sequences her vehicles based on other factors that might affect bidder participation across auctions.

A sales event i consists of J auctions; a single vehicle m is auctioned in each auction j. For notational simplicity when describing the data, we use the j subscript to refer to the auction and the vehicle offered in the auction. For each vehicle j auctioned in sales event i, the data contain an identifier for the sales event ($EventID_i$); the facility at which the sales event was conducted ($FacilityID_i$); the seller of the vehicles in the sales event ($SellerID_i$); the day the sales event was conducted (Day_i), which ranges from 1 (January 11, 2005) to 1,906 (March 31, 2010); the vehicle's model and model year ($VehicleModel_{ij}$ and $VehicleYear_{ij}$); the vehicle's run number ($RunNumber_{ij}$); the vehicle's mileage ($Mileage_{ij}$); the condition grade ($Condition_{ij}$) assigned to the vehicle, which ranges from 0 to 5 in 0.1 increments, with 0 (5) representing very poor (good) condition;[10] whether the vehicle was sold ($Sold_{ij}$); and the bid log. From the bid log, we calculated the starting bid ($StartingBid_{ij}$), the high bid ($HighBid_{ij}$), and the number of bids and bidders from (a) the physical channel ($PBids_{ij}$ and $PBidders_{ij}$), (b) the electronic channel ($EBids_{ij}$ and $EBidders_{ij}$), and (c) both channels ($Bids_{ij}$ and $Bidders_{ij}$). The data are rich, but a limitation of the data is that the number of physical bidders ($PBidders_{ij}$) is not directly recoverable from the bid log. We imputed this variable; the imputation procedure and associated implications for our estimation are discussed in the appendix. For each sold vehicle j, the data include the sales price ($SalesPrice_{ij}$), the vehicle's wholesale value as estimated by the auction company ($Valuation_{ij}$), and the ID of the winning bidder ($BidderID_{ij}$).

The electronic webcast channel was relatively new at the beginning of the sample period. During the sample period, transactions conducted electronically rose from 4.5% in 2005 to 15.7% in 2010 (see NAAA 2010, p. 8). These statistics include transactions conducted in the webcast channel as well as in other stand-alone electronic markets that operate within the wholesale used vehicle industry. Thus, they overstate the level of adoption of the electronic webcast channel that we study herein.

In the next two subsections, we examine whether use of the electronic channel led to a more even distribution of bidder participation across the sequence of auctions (§4.2) and how the evenness/unevenness of bidder participation affected seller revenue (§4.3). In other words, we examine how the electronic channel influences the demand distribution across auctions and how demand distribution, in turn, affects seller revenue.

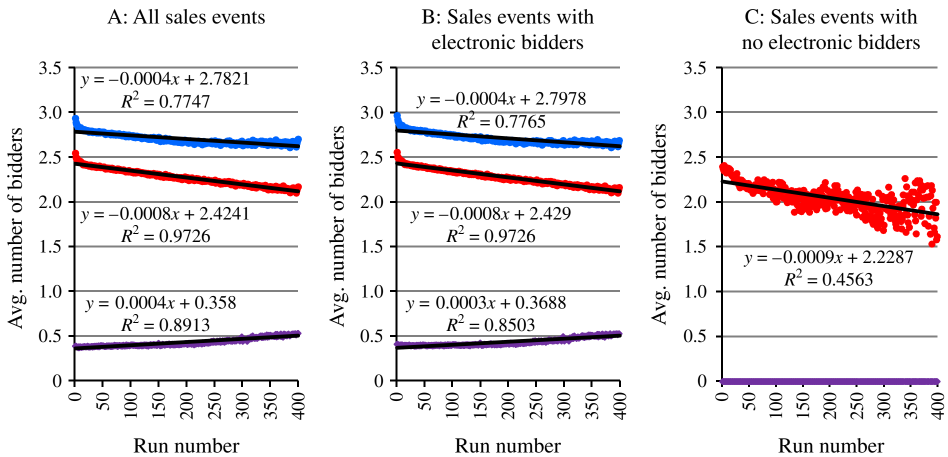
4.2. The Electronic Channel and the Distribution of Bidder Participation Across Auctions

Figure 2 plots the average number of total, physical, and electronic bidders per run number. For example, vehicles auctioned in run number 1 attracted an average of 2.93 total bidders, 2.54 of whom were physical bidders, and 0.39 of whom were electronic bidders. Figure 2 shows that the average number of physical bidders generally declines over the sequence of auctions. If electronic bidders' participation in sales events mirrored that of physical bidders, then we should see a similar pattern of attrition. However, we see the opposite; the average number of electronic bidders *increases* over the sequence. This is consistent with our argument that bidders use the electronic channel to shift their bidding from auctions with high competition (which tend to be those early in the sequence) to those with low competition (which tend to be those later in the sequence), i.e., to conduct Robin Hood operations. Electronic bidders likely behave this way because their low search costs make it easier to wait until later in the sequence to bid, and they have an incentive to do so to avoid competing with the larger number of bidders (most of

[9] We assessed whether this discontinuity affected our estimates by running our empirical models for the periods before June 2007 and after December 2008. Results are similar to those we report.

[10] $Condition_{ij}$ is null for approximately 10% of the vehicle observations. We discuss the implications of this along with the analysis.

Figure 2 (Color online) Average Number of Total, Physical, and Electronic Bidders per Run Number for All Sales Events (Panel A) and Sales Events With and Without Electronic Bidder Participation (Panels B and C)



Notes. The upper data series indicates total bidders, the middle data series indicates physical bidders, and the lower data series indicates electronic bidders. The total bidders and physical bidders data series overlap perfectly in panel C, given that the sales events summarized in panel C contain no electronic bidders. Linear regression trend lines are shown for each data series.

whom are using the physical channel) who are active at the beginning of the sequence.[11] The upward trend in electronic participation partially offsets the downward trend in physical participation, such that the average participation is more evenly distributed (on average). We conducted the analyses reported in the next three subsections to explore this further.

4.2.1. Matching Analysis. To further examine the bidder participation pattern shown in Figure 2, we estimated what the pattern would have been had the electronic channel not been available (i.e., the counterfactual pattern). This allowed us to examine whether the electronic channel led to a more even distribution of bidder participation across auctions than would otherwise occur. To do this, we leveraged the fact that, of the 65,718 sales events in the data, 4,631 had no electronic bidders participate. If the sales events with no electronic participation (referred to as "no-elec") can be used as counterfactual events for the sales events with electronic participation (referred to as "elec"), then we could attribute differences in overall participation to the electronic channel.[12]

As our first step, we created plots of the average amount of bidder participation per run number for the elec and no-elec sales events. These appear in panels B and C of Figure 2. As in panel A, the average number of physical bidders generally declines over the sequence, with the slopes of these declines similar regardless of whether the sales events had electronic bidders or not. The upward trend in the number of electronic bidders in panel B makes the overall distribution of bidder participation more even in the sales events with electronic bidding than in those without. However, if the no-elec sales events differ from the elec sales events on dimensions other than whether there was electronic participation, then the no-elec sales events would be poor counterfactual events, limiting the value of this comparison.

We used a matching procedure (both exact matching and coarsened exact matching) to address this (Iacus et al. 2011, Imbens 2004). If the no-elec sales events are appropriately matched to elec sales events, then we can be more confident that differences in bidder participation patterns are due to electronic participation as opposed to other factors. We limited our matching to sales events that occurred in 2005. The lack of electronic participation in the no-elec sales events in 2005 is likely because there was relatively low adoption of the electronic channel at this time, and bidders who did adopt did not participate in all sales events. As such, whether a sales event attracted electronic bidders was more likely to be random during this time than at later times in the sample. We matched elec sales events to no-elec sales events as follows. First, we exact matched on $FacilityID_i$ to eliminate geographic variation between the no-elec and elec sales events that might otherwise

[11] This suggests that physical bidders should continue to switch to the electronic channel until the distribution of bidder participation is perfectly even. That might occur eventually, but we do not observe that in our data, likely because the electronic channel was not sufficiently diffused during the sample period to support such an outcome.

[12] The equipment that enables the electronic channel was implemented in phases. As such, another potential approach would be to compare sales events in which the electronic channel was available to those in which it was not. This is not possible for our analysis because we require the bid log to measure bidder participation, and the bid log is only recorded for sales events in which the electronic channel was available.

confound the comparison. Second, we coarsened Day_i into three-month bins (i.e., quarters), $AvgMileage_i$ into 5,000-mile bins, $AvgValuation_i$ into \$1,000 bins, and $AvgCondition_i$ into bins of size 1. ($AvgMileage_i$ is the mean of $Mileage_{ij}$ for the J vehicles in sales event i, and $AvgValuation_i$ and $AvgCondition_i$ are analogous.) We only matched no-elec and elec sales events within the same bins, which limits variation due to time of year and the type of vehicles auctioned. Third, we matched sales events based on how the seller sequenced vehicles (e.g., from low to high value, high to low value, etc.), because the sequence could influence the auctions in which bidders participate. We did this as follows. We regressed $RunNumber_{ij}$ on $Valuation_{ij}$ for the J vehicles in each sales event i. We set $SequenceByValuation_i = 1$ if the overall regression F-statistic for sales event i was significant (at $p < 0.05$); otherwise, we set $SequenceByValuation_i = 0$.[13] We set $SequenceByValuation(\uparrow)_i = 1$ if $SequenceByValuation_i = 1$ and the relationship between $RunNumber_{ij}$ and $Valuation_{ij}$ was positive. We defined $SequenceByValuation(\downarrow)_i$ analogously. We repeated this procedure using other right-hand-side variables in place of $Valuation_{ij}$,[14] including $Condition_{ij}$, $Mileage_{ij}$, and $Supply_{ij}$ (which we measured as the count of vehicles in sales event i of the same year/model as vehicle j).[15] Summary statistics for the resulting *SequenceBy*… variables appear in the appendix. We included $SequenceByValuation(\uparrow)_i$ and $SequenceByValuation(\downarrow)_i$ in the matching criteria. This allowed us to match no-elec to elec sales events in which the seller sequenced vehicles from low to high value (i.e., $SequenceByValuation(\uparrow)_i = 1$ and $SequenceByValuation(\downarrow)_i = 0$), or from high to low value ($(\uparrow)_i = 0$ and $(\downarrow)_i = 1$), or in which the seller did not sequence vehicles by value ($(\uparrow)_i = 0$ and $(\downarrow)_i = 0$). As an example of our matching, a no-elec sales event in Dallas in the first quarter of 2005 with $AvgMileage_i$ between 40,000 and 45,000; $AvgValuation_i$ between \$10,000 and \$11,000; $AvgCondition_i$ between 3 and 4; and $SequenceByValuation(\uparrow)_i = 1$ (and therefore $SequenceByValuation(\downarrow)_i = 0$) could only be matched to an elec sales event with the same characteristics, i.e., that was in the same "cell."

The matching procedure yielded 644 cells containing 749 no-elec sales events matched to 1,006 elec sales events. Some no-elec events were matched to more than one elec event and vice versa. We refer to these sales events as the matched sample.[16] The values for Day_i, $AvgMileage_i$, and $AvgValuation_i$ are statistically indistinguishable between the no-elec and elec sales events (based on a t-test). The values for $AvgCondition_i$ are statistically different ($p < 0.01$), although the difference ($\delta = 0.04$) is small and of questionable practical significance. Overall, the matches are well balanced. Panels A and B of Figure 3 show the average number of total, physical, and electronic bidders per run number for the elec and no-elec events in the matched sample. Panels C and D show the same statistics from a variation of the matched sample in which we matched on $SequenceBySupply_i$ ($\uparrow$ and $\downarrow$) rather than $SequenceByValuation_i$ ($\uparrow$ and $\downarrow$). The patterns are similar when we match on $SequenceByMileage_i$ ($\uparrow$ and $\downarrow$) and $SequenceByCondition_i$ ($\uparrow$ and $\downarrow$). We did not match on all eight of the *SequenceBy*… variables simultaneously because that created substantial dimensionality in the matching procedure and reduced the size of the matched sample too greatly. (We control for all of the sequencing variables simultaneously in §4.2.2.) Overall, Figure 3 shows that when electronic bidders participate (a) the distribution of bidders across auctions is more even, and (b) more bidders participate (i.e., we see both demand distribution and demand expansion). The evidence of demand distribution is consistent with our hypothesis that bidders leverage the lower search costs of the electronic channel to shift participation from high to low competition auctions, which results in a more even distribution of bidder participation.

4.2.2. Vehicle-Level Regressions. We explored further whether the trends shown in Figures 2 and 3 might be a result of the types of vehicles offered in each run number, rather than a result of differences in search costs across the channels (as we conclude). For example, prior research has shown that, ceteris paribus, electronic bidders prefer relatively low-mileage, high-value vehicles whose quality is predictable and that can be easily represented online (Overby and Jap 2009). If these vehicles tend to be auctioned later in the sequence, then that could

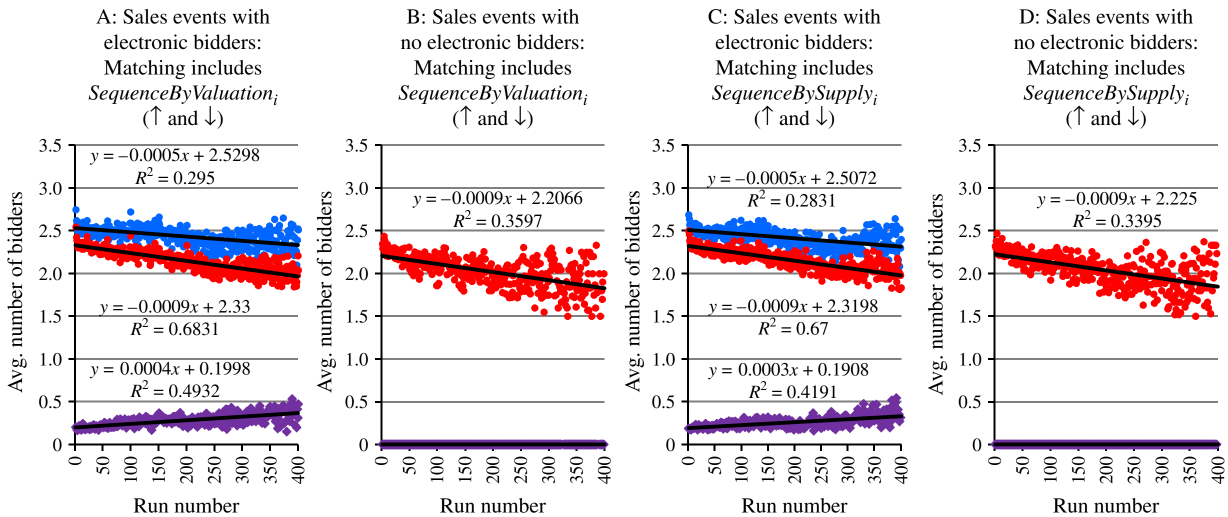
[13] If the F-statistic was insignificant for a sales event (perhaps because $RunNumber_{ij}$ and $Valuation_{ij}$ are related, but not in a linear fashion), then we rechecked the F-statistic after rerunning the regression with $Valuation_{ij}^2$ (and then $Valuation_{ij}^2$ and $Valuation_{ij}^3$) added to the right-hand side.

[14] An alternative approach is to include all right-hand-side variables in the same regression for each sales event i. Both approaches allow us to assess whether multiple variables are related to $RunNumber_{ij}$ in a given sales event. We judged our approach to have more power to detect relationships because each variable is examined individually.

[15] We used $Supply_{ij}$ because Grether and Plott (2009) noted that many sellers in this industry sequence vehicles by supply. They also noted that the best sequencing strategy is one that is opposite to that which other sellers are using, i.e., there is no dominant strategy, which explains the heterogeneity in vehicle sequencing (see the appendix).

[16] The matched sample does not contain sales events that could not be matched. We report the data for all no-elec and elec sales events in Figure 2 to show that the results of the matched sample are consistent with the full sample.

Figure 3 (Color online) Average Number of Total, Physical, and Electronic Bidders per Run Number for Sales Events With and Without Electronic Bidder Participation in the Matched Sample



Notes. The upper data series indicates total bidders, the middle data series indicates physical bidders, and the lower data series indicates electronic bidders. The total bidders and physical bidders data series overlap perfectly in panels B and D, given that the sales events summarized in panels B and D contain no electronic bidders. Linear regression trend lines are shown for each data series.

explain the trends. We controlled for this in the matching estimation by matching on the *SequenceBy…* variables, and we explored this further via the following regression model:

$$
\begin{aligned}
DV_{ij} = {} & \beta_0 + \beta_1 * RunNumber_{ij} + \beta_2 * Mileage_{ij} \\
& + \beta_3 * Condition_{ij} + \beta_4 * Valuation_{ij} \\
& + \beta_5 * ValuationMinusStartingBid_{ij} + \beta_6 * Supply_{ij} \\
& + \beta_7 * VehiclesOffered_i + \beta_8 * AvgMileage_i \\
& + \beta_9 * AvgCondition_i + \beta_{10} * AvgValuation_i \\
& + \beta_{11} * Day_i + \sum_{k=2}^{K} \beta_{12(k)} * Seller(k)_i + \varepsilon_{ij}. \qquad (2)
\end{aligned}
$$

In (2), DV_{ij} is $Bidders_{ij}$, $PBidders_{ij}$, or $EBidders_{ij}$. Note that $Bidders_{ij} = PBidders_{ij} + EBidders_{ij}$. Using $Bidders_{ij}$ as the dependent variable allows us to examine the overall relationship between run number and the number of bidders after controlling for vehicle quality (e.g., mileage, valuation), a time trend, seller fixed effects, and other factors. Using $PBidders_{ij}$ and $EBidders_{ij}$ as the dependent variables allows us to decompose the overall effect to see how the relationship between run number and the number of bidders differs based on bidder type. In (2), $Seller(k)_i$ are indicator variables for each seller k (i.e., fixed effects), and the other variables are described in Table 1. Notice that (2) contains variables at both the vehicle and sales event levels. This is because the 10,351,857 vehicles in the data are clustered within the 65,718 sales events. We accounted for the clustered nature of the data in two ways: (a) by clustering the standard errors by sales event (this accounts for correlation among the error terms for vehicles in the same sales event), and (b) by estimating a multilevel model (aka a hierarchical linear model), the specification for which is shown as (3). In the multilevel model, we allow the intercept to vary by sales event (note the i subscript in β_{0i} in (3)), and we model this variance as a function of the sales event variables. The number of observations and explanatory variables (particularly given the inclusion of the seller fixed effects) increases model dimensionality substantially, making convergence of the multilevel model difficult. Accordingly, we fit this model using a random 10% sample of the sales events. We used the full data when fitting model (2):

$$
\begin{aligned}
DV_{ij} = {} & \beta_{0i} + \beta_1 * RunNumber_{ij} + \beta_2 * Mileage_{ij} \\
& + \beta_3 * Condition_{ij} + \beta_4 * Valuation_{ij} \\
& + \beta_5 * ValuationMinusStartingBid_{ij} \\
& + \beta_6 * Supply_{ij} + r_{ij}, \\
\text{where } \beta_{0i} = {} & \gamma_{00} + \gamma_{01} * VehiclesOffered_i \\
& + \gamma_{02} * AvgMileage_i + \gamma_{03} * AvgCondition_i \\
& + \gamma_{04} * AvgValuation_i + \gamma_{05} * Day_i \\
& + \sum_{k=2}^{K} \gamma_{06(k)} * Seller(k)_i + u_{0i}. \qquad (3)
\end{aligned}
$$

We refer to these as the vehicle-level regressions. Descriptive statistics appear in Table 1, and results using specification (2) appear in Table 2 (results from (3) are similar and are available from the authors).

Table 1 Descriptive Statistics for Variables in the Vehicle-Level Regressions

Variable	Description	Mean	SD	Min	Max	Median
$Bidders_{ij}$	Number of bidders for vehicle j	2.75	1.28	0	26	2
$PBidders_{ij}$	Number of physical bidders for vehicle j	2.33	1.04	0	20	2
$EBidders_{ij}$	Number of electronic bidders for vehicle j	0.42	0.80	0	20	0
$RunNumber_{ij}$	Run number in which vehicle j is offered	171.75	164.07	1	802	127
$Mileage_{ij}{}^{a}$	Mileage of vehicle j	4.40	3.83	0.0010	39.83	3.12
$Condition_{ij}$	Condition grade of vehicle j	3.07	0.83	0	5	3
$Valuation_{ij}{}^{a}$	Estimated value of vehicle j	1.29	0.73	0.0025	35.5	1.21
$ValuationMinus\ StartingBid_{ij}{}^{b}$	$Valuation_{ij}$ minus $StartingBid_{ij}$ for vehicle j	1.22	1.42	−4.9	9.9	0.97
$Supply_{ij}{}^{c}$	Number of vehicles of same year/model as vehicle j (includes vehicle j)	1.48	2.65	0.1	65.1	0.40
$VehiclesOffered_{i}{}^{d}$	Number of vehicles offered in sales event i	2.40	1.20	0.01	8.02	2.29
$AvgMileage_{i}{}^{a}$	Average mileage of vehicles in sales event i	4.43	2.63	0.0102	28.85	3.83
$AvgCondition_{i}$	Average condition of vehicles in sales event i	3.06	0.50	0	5	3.05
$AvgValuation_{i}{}^{a}$	Average valuation of vehicles in sales event i	1.28	0.48	0.0225	20.17	1.27
Day_{i}	Day sales event occurred (January 11, 2005 = 1)	908.53	629.03	1	1,906	718

Correlations	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. $Bidders_{ij}$	1													
2. $PBidders_{ij}$	0.72	1												
3. $EBidders_{ij}$	0.58	−0.06	1											
4. $RunNumber_{ij}$	−0.01	−0.06	0.05	1										
5. $Mileage_{ij}{}^{a}$	−0.03	0.07	−0.13	−0.10	1									
6. $Condition_{ij}$	0.06	−0.04	0.12	0.05	−0.38	1								
7. $Valuation_{ij}{}^{a}$	0.11	−0.06	0.23	0.07	−0.52	0.30	1							
8. $ValuationMinusStartingBid_{ij}{}^{b}$	0.16	0.14	0.08	−0.04	0.09	−0.27	0.18	1						
9. $Supply_{ij}{}^{c}$	−0.05	−0.12	0.07	0.20	−0.28	0.14	0.13	−0.13	1					
10. $VehiclesOffered_{i}{}^{d}$	0.01	−0.05	0.10	0.38	−0.26	0.19	0.19	−0.14	0.38	1				
11. $AvgMileage_{i}{}^{a}$	−0.01	0.11	−0.14	−0.20	0.61	−0.34	−0.43	0.19	−0.37	−0.43	1			
12. $AvgCondition_{i}$	0.01	−0.08	0.10	0.11	−0.35	0.60	0.29	−0.26	0.24	0.32	−0.58	1		
13. $AvgValuation_{i}{}^{a}$	0.07	−0.09	0.22	0.14	−0.40	0.25	0.66	0.02	0.28	0.28	−0.65	0.43	1	
14. Day_{i}	0.09	−0.02	0.12	−0.01	0.08	0.04	−0.00	0.02	−0.11	−0.04	0.14	0.04	0.01	1

[a]Scaled by dividing by 10,000.
[b]Scaled by dividing by 1,000.
[c]Scaled by dividing by 10.
[d]Scaled by dividing by 100.

The largest variance inflation factor (VIF) is 3.58, suggesting that multicollinearity is not a major concern. We also estimated (2) using other specifications, including negative binomial, zero-inflated negative binomial, and seemingly unrelated regression (to estimate the model for all three dependent variables simultaneously).[17] Results are similar to those in Table 2. Results indicate that the relationships shown in Figures 2 and 3 hold after controlling for other variables/alternative explanations. Specifically, there is a negative relationship between $RunNumber_{ij}$ and $Bidders_{ij}$, a (more) negative relationship between $RunNumber_{ij}$ and $PBidders_{ij}$, and a positive relationship between $RunNumber_{ij}$ and $EBidders_{ij}$.

4.2.3. Range of Purchasing Behavior. In §3, we argued that lower search costs allow electronic bidders to observe and potentially participate in more auctions, thereby helping them to conduct Robin Hood operations that might lead to the participation patterns shown in Figure 2. We examined this by investigating whether bidders using the electronic channel remained active in sales events longer than bidders using the physical channel. We did this by examining instances ($n = 1{,}379{,}379$) in which a given bidder (based on his *BidderID*) purchased at least two vehicles in the same sales event i. For each of these instances, we recorded the lowest and highest run number at which the bidder purchased and subtracted them to approximate the range of the bidder's

[17] We used a negative binomial model because the dependent variables are counts that are overdispersed. We used a zero-inflated negative binomial model because the dependent variables, particularly $EBidders_{ij}$, are often 0. We did this purely for robustness, however, because a zero-inflated model assumes a distinct process by which some of the observations must always equal 0 (Cameron and Trivedi 2005), and there is no such process in our context (i.e., there is no structural reason why some vehicles would always attract 0 bidders). The seemingly unrelated regression (SUR) results are identical to those reported in Table 2 because SUR yields the same results as estimating each regression separately when the right-hand-side variables are the same (see Wooldridge 2002, Theorem 7.6, p. 164).

Table 2 Results of the Vehicle-Level Regressions

	Dep. var.: $Bidders_{ij}$ Coefficient	Dep. var.: $PBidders_{ij}$ Coefficient	Dep. var.: $EBidders_{ij}$ Coefficient
$RunNumber_{ij}$	−0.0001 (0.0000)***	−0.0002 (0.0000)***	0.0001 (0.0000)***
$Mileage_{ij}$[a]	0.0042 (0.0003)***	0.0046 (0.0002)***	−0.0004 (0.0002)*
$Condition_{ij}$	0.1598 (0.0013)***	0.0471 (0.0008)***	0.1128 (0.0010)***
$Valuation_{ij}$[a]	0.0531 (0.0022)***	−0.0178 (0.0014)***	0.0709 (0.0017)***
$ValuationMinusStartingBid_{ij}$[b]	0.1380 (0.0016)***	0.0747 (0.0009)***	0.0634 (0.0009)***
$Supply_{ij}$[c]	−0.0277 (0.0009)***	−0.0214 (0.0006)***	−0.0063 (0.0006)***
$VehiclesOffered_i$[d]	0.0297 (0.0021)***	0.0109 (0.0013)***	0.0188 (0.0019)***
$AvgMileage_i$[a]	0.0019 (0.0016)	0.0154 (0.0012)***	−0.0135 (0.0015)***
$AvgCondition_i$	−0.0893 (0.0062)***	−0.0675 (0.0041)***	−0.0218 (0.0062)***
$AvgValuation_i$[a]	0.0390 (0.0116)***	−0.0816 (0.0070)***	0.1206 (0.0104)***
Day_i	0.0002 (0.0000)***	−0.0001 (0.0000)***	0.0003 (0.0000)***
Intercept[e]	2.2974 (0.0510)***	2.6272 (0.0423)***	−0.3298 (0.0426)***
Seller fixed effects[e]	included	included	included
n[f]	6,530,431	6,530,431	6,530,431
R^2	0.08	0.06	0.17

Notes. Results are from specification (2) in the text in which the standard errors are clustered by sales event. Results of the multilevel model shown in specification (3) are available from the authors.

[a]Variable scaled by dividing by 10,000.

[b]Variable scaled by dividing by 1,000.

[c]Variable scaled by dividing by 10.

[d]Variable scaled by dividing by 100.

[e]The overall intercept is reported but has minimal meaning because it represents the conditional mean of the dependent variable for the seller whose fixed effect was withheld.

[f]Including $Valuation_{ij}$ and $Condition_{ij}$ reduces the sample size because $Valuation_{ij}$ is only recorded for vehicles that sold, and $Condition_{ij}$ is null for approximately 10% of the transactions. Similar results are achieved after dropping $Valuation_{ij}$, $Condition_{ij}$, or both.

$^{*}p < 0.10$; $^{***}p < 0.01$.

participation in the sales event. For example, if a bidder used the physical channel to purchase vehicles at run numbers 28, 49, and 90, then his range of participation was $90 - 28 = 62$. The average range was 88.7 (SD = 88.0) and 101.4 (SD = 91.8) for physical and electronic bidders, respectively. The range was significantly higher ($p < 0.01$) for electronic bidders than for physical bidders, based on either a t-test or a Mann–Whitney U test. A limitation of this measure is that it is based on auctions won rather than auctions observed. Thus, it may underestimate the true range of a buyer's participation. However, this issue will exist in both channels, such that a comparison between channels should still yield useful information about differences in participation.

4.3. The Distribution of Bidder Participation Across Auctions and Seller Revenue

We next examined how the evenness/unevenness of the distribution of bidder participation across the auctions in a sales event affects seller revenue. Our empirical analysis is motivated directly by the analytical model. The analytical model considers how demand distribution across a group of single-object auctions affects seller revenue. Because each sales event consists of a group of single-object auctions, we used the sales event i as our unit of analysis. The key variables identified by the analytical model are (a) expected revenue from the auctions, (b) the distribution of bidder participation across auctions, (c) the average number of bidders per auction, (d) the degree to which the variability of bidder valuations varies across the objects being auctioned, and (e) whether the seller uses knowledge of the distribution of bidder valuations to allocate objects to auctions. We measured these as follows.

Expected Revenue: The analytical model considers total expected revenue across a group of auctions, although adjusting the model to consider average expected revenue for each auctioned object is trivial. To make it easier to interpret the economic significance of our empirical estimates, we measured average expected revenue for each vehicle auctioned in sales event i. Specifically, $AvgHighBid_i = \sum_{j=1}^{J} HighBid_{ij}/J$. Because $HighBid_{ij}$ is equal in expectation to the second-highest order statistic from the set of bidder valuations, it corresponds to the expected auction price we use in the analytical model.

Distribution of Bidder Participation Across Auctions: We measured the distribution of bidder participation across vehicles in each sales event ($StDevBidders_i$) as the standard deviation of $Bidders_{ij}$. For example, if $Bidders_{ij}$ is the same for all vehicles in a sales event, then $StDevBidders_i = 0$. Thus, a lower value of $StDevBidders_i$ indicates a more even distribution of bidder participation. This relates to the Robin Hood operations described in §3 because each Robin Hood operation will reduce $StDevBidders_i$.[18]

Average Number of Bidders per Auction: We measured the average number of bidders for a vehicle j in sales event i as $AvgBidders_i = \sum_{j=1}^{J} Bidders_{ij}/J$.

Degree to Which the Variability of Bidder Valuations Varies Across the Objects Being Auctioned: We measured this via two steps. First, we estimated the variability of bidder valuations for each vehicle j ($ValuationStDev\_Vehicle_{jt}$) as the standard deviation of the high bid for all vehicles of the same year and model as vehicle j auctioned in the same month t. (We also computed the corresponding mean, $ValuationMean\_Vehicle_{jt}$.) Note that $ValuationStDev\_Vehicle_{jt}$ measures the variability of the second-highest order statistic (i.e., the auction price) for vehicles of a given year/model rather than the variability of the underlying bidder valuations. However, this is a reasonable approach for our purposes because a low (high) standard deviation of high bids should reflect relatively low (high) variability.[19] Second, we calculated the degree to which the variability of bidder valuations for each vehicle varies across a sales event i ($ValuationStDev\_SalesEvent_i$) as the standard deviation of $ValuationStDev\_Vehicle_{jt}$ for the vehicles auctioned in sales event i. For example, if an entire sales event consisted of 2007 Ford Rangers (with $ValuationStDev\_Vehicle_{jt} = 1{,}500$), then $ValuationStDev\_SalesEvent_i = 0$. If the sales event consisted of five 2007 Ford Rangers and three 2007 Ford Explorers (with $ValuationStDev\_Vehicle_{jt} = 2{,}500$), then $ValuationStDev\_SalesEvent_i = 517.55$.

Whether the Seller Uses Knowledge of the Distribution of Bidder Valuations to Allocate Objects to Auctions: If the seller in sales event i sequences her vehicles (i.e., allocates her objects to auctions) based on knowledge of the distribution of bidder valuations, then we would expect a correlation between $RunNumber_{ij}$ and $ValuationStDev\_Vehicle_{jt}$. We examined this by regressing $RunNumber_{ij}$ on $ValuationStDev\_Vehicle_{jt}$, $ValuationStDev\_Vehicle^2_{jt}$, and $ValuationStDev\_Vehicle^3_{jt}$ for the J vehicles in each sales event i. This allows a potential relationship between $RunNumber_{ij}$ and $ValuationStDev\_Vehicle_{jt}$ in sales event i to take any form that can be approximated by a third-degree polynomial. We set $SequenceByValDist_i = 1$ if the overall regression F-statistic was significant (at $p < 0.05$). As noted in §3.2, this is the relevant sequencing method for the analytical model. Other sequencing methods such as by $Mileage_{ij}$ or by $Valuation_{ij}$ are not relevant, unless they are surrogates for sequencing based on the bidder valuation distributions (which will be picked up by $SequenceByValDist_i$).

4.3.1. Regression Specification. To examine the relationships posited by the analytical model, we regressed $AvgHighBid_i$ on $StDevBidders_i$, $AvgBidders_i$, $ValuationStDev\_SalesEvent_i$, $SequenceByValDist_i$, interactions between $StDevBidders_i$ and the other aforementioned variables, and a set of control variables. We refer to this as the sales event regression. Including $AvgBidders_i$ in the regression allowed us to control for any demand expansion effect so that we can better identify the demand distribution effect represented by $StDevBidders_i$. We include the interaction terms because the analytical model shows that the relationship between the distribution of bidder participation across vehicles ($StDevBidders_i$) and expected revenue ($AvgHighBid_i$) should be (a) weaker at higher levels of $AvgBidders_i$, because a more even distribution has less impact on revenue when the number of bidders is high; (b) weaker at higher levels of $ValuationStDev\_SalesEvent_i$, because a more even distribution is less likely to increase revenue when bidder valuations vary widely across objects; and (c) weaker if $SequenceByValDist_i = 1$, because a more even distribution is less likely to increase revenue when the seller allocates objects to auctions based on (a priori) knowledge of the distribution of bidder valuations. We mean-centered the variables that appear in the interaction terms so that the coefficients on the lower-order terms represent main effects rather than simple effects (Echambadi and Hess 2007).

[18] This measure reflects how we measure the evenness of the distribution of bidder participation in the analytical model because if vector X_1 majorizes vector X_2, then the standard deviation of the elements in vector X_1 will exceed that of the elements in vector X_2 (Arnold 1987). However, the converse is not always true, i.e., if the standard deviation of the elements in vector X_1 exceeds that of the elements in vector X_2, then X_1 does not always majorize X_2. In the case when neither vector majorizes the other, our theoretical model does not predict which distribution of bidder participation yields higher revenues. To assess how well our empirical measure corresponds to the underlying theoretical construct, we simulated how often a more evenly distributed vector of bidder participation yields higher revenue than a less evenly distributed vector (as indicated by the standard deviation of bidders across auctions), even if the latter does not majorize the former. In the simulation, we generated two vectors of bidder participation (X_1 and X_2) across $J = 158$ auctions (which is the mean in our empirical application). The distribution of bidder participation differed between the two vectors, but the aggregate amount did not. We calculated the total expected revenue from the J auctions for each scenario by assuming that bidder valuations for each object were uniformly distributed between 10,000 and 12,000. We ran the simulation 1,000 times. The vector with the lower standard deviation yielded higher total expected revenue 93% of the time, with the increase equal to approximately \$14.50 per vehicle. For the exceptions in which the higher standard deviation vector yielded higher total expected revenue, the increase was approximately \$2.35 per vehicle. This indicates that $StDevBidders_i$ is a good empirical measure of our theoretical construct.

[19] Because we calculated it for each vehicle year/model on a monthly basis, $ValuationStDev\_Vehicle_{jt}$ is not artificially inflated by differences in vehicle age or model. Although our procedure does not explicitly control for differences in vehicle mileage and condition, these variables are highly correlated with vehicle age.

Control variables include the vehicle quality measures $AvgMileage_i$, $AvgValuation_i$, and $AvgCondition_i$;[20] $VehiclesOffered_i$; Day_i (to control for a linear time trend); and seller fixed effects (to control for unobserved seller-specific factors). Descriptive statistics are shown in Table 3. Results are shown in Table 4. All of the VIFs for the independent variables in the model without the interaction terms are 2.20 or below. This suggests that collinearity is unlikely to be a problem for our estimates.[21]

4.3.2. Results. The coefficient for $StDevBidders_i$ is negative and significant, indicating that seller revenue is lower for sales events with less even distributions of bidders across auctions, which is consistent with the analytical model. The coefficient for $AvgBidders_i$ is positive and significant. This is consistent with auction theory that more bidders are associated with higher prices. To account for a possible nonlinear relationship between $AvgBidders_i$ and $AvgHighBid_i$, we reestimated the model with both $AvgBidders_i$ and $AvgBidders_i^2$. However, neither coefficient was significant when both were included, so we retained the model with $AvgBidders_i$ only.[22] The coefficient for $StDevBidders_i * AvgBidders_i$ is positive and significant, indicating that the negative effect of $StDevBidders_i$ is weaker at higher levels of $AvgBidders_i$. This is consistent with the analytical model. The coefficient for $StDevBidders_i * SequenceByValDist_i$ is positive and significant, indicating that the negative effect of $StDevBidders_i$ is attenuated when the seller sequences the vehicles based on the variability of the bidder valuations. This is also consistent with the analytical model. The coefficient for $StDevBidders_i * ValuationStDev\_SalesEvent_i$ is positive (which is consistent with the analytical model) but insignificant.

4.3.3. Economic Significance. To get a sense of the economic significance of these results to sellers in the wholesale used vehicle market, we considered the marginal effect of a change of one standard deviation in $StDevBidders_i$ (0.27; see Table 3) when $AvgBidders_i$ is at its mean (2.70) and $SequenceByValDist_i = 0$. This effect is -67.56, ceteris paribus. For a sales event of average length (i.e., 158 vehicles), a \$67.56 decrease in the average high bid equates to a reduction of \$10,675 in total revenue for the entire sales event, assuming that the seller accepts all of the high bids. In the appendix, we show that this estimate may be conservative because of possible measurement error in $StDevBidders_i$ and $AvgBidders_i$. The most powerful explanatory variables are $AvgValuation_i$ and $AvgCondition_i$ (based on the standardized coefficients shown in Table 4), which is intuitive given that the average high bid for vehicles depends heavily on the value and condition of the vehicles.

4.3.4. Examining the Regression Specification/ Robustness Checks. We considered whether the coefficients reported in Table 4 could be biased because of specification error. A particular problem in the analysis of auctions is endogenous entry (e.g., Bajari and Hortacsu 2003). This problem arises because the number of bidders in an auction is not fixed; rather, bidders choose whether to participate. If unobserved

[20] Although $Valuation_{ij}$ and $Condition_{ij}$ are null for some vehicles, we are able to calculate $AvgValuation_i$ and $AvgCondition_i$ using the vehicles in a sales event for which we observe these values. We believe these to be effective controls for the average quality of the vehicles in sales event i. However, if the missing values of $Valuation_{ij}$ and $Condition_{ij}$ are systematically different from the observed values, then these controls will not work as intended. To examine this, we imputed the missing values. We imputed $Valuation_{ij}$ using a procedure similar to that which the auction company uses to calculate $Valuation_{ij}$. We grouped vehicles by year and model. We took the sold transactions for each group in a given calendar year and regressed $HighBid_{ij}$ on $Mileage_{ij}$. We used the resulting coefficients, including the intercept, to impute $Valuation_{ij}$ for each vehicle of that year and model (and mileage) that was offered in that calendar year; we label this $Valuation\_Imputed_{ij}$. We assessed the performance of this procedure by regressing the observed values of $Valuation_{ij}$ on $Valuation\_Imputed_{ij}$. The coefficient for $Valuation\_Imputed_{ij}$ is 0.99 ($p < 0.01$), and $R^2 = 0.95$. This indicates that our imputation procedure closely reflects $Valuation_{ij}$ as estimated by the auction company. We imputed $Condition_{ij}$ as follows. For each vehicle of year/model j offered by seller k on day t, we calculated the average condition grade for vehicles of the same year/model j offered by the same seller k over the prior 30 days ($AvgCondition_{jkt-30}$). We regressed $Condition_{ij}$ (when observed) on $AvgCondition_{jkt-30}$, $Mileage_{ij}$, and the vehicle's age (computed as the date of the sales event minus January 1 of $VehicleYear_{ij}$). We used the resulting coefficients to impute $Condition_{ij}$, which we label $Condition\_Imputed_{ij}$. As we did for $Valuation_{ij}$, we regressed the observed values of $Condition_{ij}$ on $Condition\_Imputed_{ij}$. The coefficient for $Condition\_Imputed_{ij}$ is 0.99 ($p < 0.01$), and $R^2 = 0.49$. We then replaced the missing values of $Valuation_{ij}$ and $Condition_{ij}$ with the imputed values and recalculated $AvgValuation_i$ and $AvgCondition_i$. The results of the sales event regression are qualitatively the same whether we use the versions of $AvgValuation_i$ and $AvgCondition_i$ with the imputed values or those without.

[21] We also computed the VIFs for the interaction terms using the *uncentered* data as suggested by Echambadi and Hess (2007). VIFs for uncentered interaction terms are often high, because main effects are typically highly correlated with the interaction effects of which they are a part (and vice versa). Ours are no different, with VIF = 49.2 for the *uncentered* $StDevBidders_i * AvgBidders_i$ interaction term. This is an artifact of the definition of the VIF rather than a threat to valid estimation, as discussed by Allison (2012; archived by WebCite® at http://www.webcitation.org/6R7G70Aou). All VIFs for the *mean-centered* interaction terms are below 2.

[22] The lack of significance for the linear and quadratic terms when included together may be because this model is at the sales event level, i.e., at the aggregate level. To examine this, we ran a model at the *vehicle level* in which we regressed $HighBid_{ij}$ on a constant, $Bidders_{ij}$, $Bidders_{ij}^2$, $Valuation_{ij}$, $Condition_{ij}$, and $Mileage_{ij}$. The coefficient for $Bidders_{ij}$ is 205.47 (SE = 2.18), and the coefficient for $Bidders_{ij}^2$ is -6.39 (SE = 0.28). This reflects the concave relationship between $HighBid_{ij}$ and $Bidders_{ij}$ predicted by theory.

Table 3 Descriptive Statistics for Variables in the Sales Event Regression

Variable	Description (each variable describes a sales event i)	Mean	SD	Min	Max	Median
$AvgHighBid_i$	Average high bid	117,84	5,721	132	148,299	11,745
$StDevBidders_i$	Standard deviation of the number of bidders	1.19	0.27	0	3.28	1.18
$AvgBidders_i$	Average number of bidders	2.70	0.48	0.02	6.4	2.68
$ValuationStDev\_SalesEvent_i$	Standard deviation of the standard deviation of bidder valuations for vehicles in sales event i	1,069	1,630	0	10,780	883.97
$SequenceByValDist_i$	Indicator for whether the seller sequenced vehicles based on the bidder valuation distributions	0.39	0.49	0	1	0
$AvgValuation_i$	Average vehicle valuation	12,126	5,439	225	201,666	11,616
$AvgMileage_i$	Average vehicle mileage	52,411	32,179	102	288,545	45,007
$AvgCondition_i$	Average vehicle condition	2.93	0.58	0	5	3
$VehiclesOffered_i$	Number of vehicles offered	157.52	113.94	10	802	139
Day_i	Day sales event occurred (January 11, 2005 = 1)	891.24	634.24	1	1,906	703

Correlations	1	2	3	4	5	6	7	8	9	10
1. $AvgHighBid_i$	1									
2. $StDevBidders_i$	−0.05	1								
3. $AvgBidders_i$	0.11	0.62	1							
4. $ValuationStDev\_SalesEvent_i$	0.05	0.05	0.04	1						
5. $SequenceByValDist_i$	0.02	0.01	−0.00	−0.00	1					
6. $AvgValuation_i$	0.97	−0.03	0.10	0.06	0.01	1				
7. $AvgMileage_i$	−0.65	0.17	0.05	0.04	−0.03	−0.64	1			
8. $AvgCondition_i$	0.54	−0.17	−0.04	−0.03	0.04	0.46	−0.55	1		
9. $VehiclesOffered_i$	0.26	0.01	0.09	0.00	0.23	0.22	−0.34	0.30	1	
10. Day_i	0.01	0.11	0.20	0.08	0.04	0.01	0.11	0.04	0.03	1

Notes. Additional detail about variable definitions is available in the text. The descriptive statistics for $AvgValuation_i$, $AvgMileage_i$, $AvgCondition_i$, and $VehiclesOffered_i$ differ from those reported in Table 1 because Table 1 summarizes data at the vehicle level; here, data are summarized at the sales event level. For example, suppose there were only two sales events: A with two vehicles and $AvgMileage_i = 40{,}000$ and B with four vehicles and $AvgMileage_i = 60{,}000$. The mean of $AvgMileage_i$ would be 53,333 in Table 1 and 50,000 here.

factors influence both bidder participation and the auction price, then bidder participation variables will be endogenous in a regression on prices (this can be thought of as either an omitted variables problem or a simultaneity problem). For example, unobserved item quality and/or unobserved seller characteristics might create endogeneity in a regression model such as ours. We believe that endogeneity arising from factors such as these is unlikely in our case, given that we have controlled for vehicle quality

Table 4 Results of the Sales Event Regression

	Dep. var.: $AvgHighBid_i$		Dep. var.: $AvgHighBid_i$	
	Coef. (robust SE)	Std. coef.[a]	Coef. (robust SE)	Std. coef.[a]
$AvgValuation_i$	0.927 (0.027)***	0.880	0.927 (0.027)***	0.880
$AvgMileage_i$	−0.001 (0.002)	−0.004	−0.001 (0.002)	−0.004
$AvgCondition_i$	1,101.773 (33.275)***	0.111	1,102.342 (33.206)***	0.111
$VehiclesOffered_i$	1.111 (0.071)***	0.022	1.136 (0.073)***	0.023
Day_i	−0.102 (0.010)***	−0.011	−0.105 (0.010)***	−0.012
$StDevBidders_i$	−538.596 (48.615)***	−0.026	−531.553 (48.214)***	−0.025
$AvgBidders_i$	396.658 (40.748)***	0.033	384.162 (40.968)***	0.032
$ValuationStDev\_SalesEvent_i$	0.017 (0.011)	0.005	0.018 (0.011)	0.005
$SequenceByValDist_i$	10.084 (8.867)	0.001	10.635 (8.873)	0.001
$StDevBidders_i * AvgBidders_i$			104.19 (52.49)**	0.003
$StDevBidders_i * ValuationStDev\_SalesEvent_i$			0.003 (0.044)	0.000
$StDevBidders_i * SequenceByValDist_i$			136.239 (41.970)***	0.003
Intercept[b]	-35,947.4 (1,466.11)***		-35,626.2 (1,469.79)***	
Seller fixed effects[b]	Included		Included	
n	65,718		65,718	
R^2	0.97		0.97	

[a]These are standardized regression coefficients.

[b]The overall intercept is reported but has minimal meaning because it represents the conditional mean of the dependent variable for the seller whose fixed effect was withheld.

$^{**}p < 0.05$; $^{***}p < 0.01$.

via $AvgValuation_i$, $AvgMileage_i$, and $AvgCondition_i$; we have controlled for seller characteristics via seller fixed effects; and we have included a linear time trend via Day_i to control for general unobserved factors that vary over time. A related potential problem is if high prices cause lower bidder participation by filtering out low-valuation bidders who would have otherwise participated. This might create reverse causality between $AvgHighBid_i$ and the bidder participation variables. However, this is not actually a problem for our model (at least not directly), because it is not prices (i.e., $AvgHighBid_i$) that affect bidder participation, it is where the auctioneer sets the starting bid. For example, consider two vehicles that sell for $5,000 and $30,000, respectively. If the auctioneer starts the bidding for the $5,000 ($30,000) vehicle at $4,900 ($29,900), then few bidders will bid. If he starts the bidding at $3,000 ($20,000), then many bidders will bid.

As a robustness check, we considered whether where the auctioneer set the starting bid could affect our results. We estimated where the auctioneer started the bidding for each vehicle relative to estimated bidder valuations for that vehicle as follows. We estimated the low and high ends of the bidder valuation distribution for vehicle j ($ValuationLow\_Vehicle_{jt}$ and $ValuationHigh\_Vehicle_{jt}$) as $ValuationMean\_Vehicle_{jt} \pm 2^* \; ValuationStDev\_Vehicle_{jt}$ (the latter two variables are defined in §4.3). We set $RelativeStartingBid_{ij} = (StartingBid_{ij} - ValuationLow\_Vehicle_{jt})/(ValuationHigh\_Vehicle_{jt} - ValuationLow\_Vehicle_{jt})$. For example, if $StartingBid_{ij} = 9{,}000$, $ValuationMean\_Vehicle_{jt} = 10{,}000$, and $ValuationStDev\_Vehicle_{jt} = 1{,}000$ (such that $ValuationLow\_Vehicle_{jt} = 8{,}000$ and $ValuationHigh\_Vehicle_{jt} = 12{,}000$), then $RelativeStartingBid_{ij} = 0.25$. We then calculated $AvgRelativeStartingBid_i$ as the mean of $RelativeStartingBid_{ij}$ for the J vehicles in sales event i. We reestimated the sales event regressions using only those sales events for which $AvgRelativeStartingBid_i$ was in the lowest quartile (0.23 or below), i.e., those in which the auctioneer set the starting bids relatively low (on average). Results (available from the authors) are substantively similar to those in Table 4. This provides evidence that our main results are robust to the possibility that our measures of the number of bidders are truncated by high auctioneer starting bids for some sales events.

Overall, the regression model maps directly to the theory and causality developed in the analytical model. The high level of correspondence between the results of the regression model and the predictions of the analytical model (which include not only the main effect of $StDevBidders_i$ but also the interaction effects with $AvgBidders_i$ and $SequenceByValDist_i$) provides evidence that the regression model is an appropriate representation of how the theory applies to the data.

5. Conclusion

5.1. Contributions

As advances in information technology continue to reduce search costs and improve bidders' ability to participate in auctions regardless of location, bidder participation patterns will change in interesting ways. Analyzing these changes is important for extending auction theory and for understanding how sellers can leverage multiple-object auctions to maximum effect. From a theoretical perspective, we identified two effects of the Internet and reduced search costs on prices in a multiple-object auction setting. Because bidding electronically reduces search costs relative to bidding physically, we argue that this reduction in search costs may cause bidders to participate not only in more auctions (the "demand expansion" effect) but also to adjust their participation so that bidding competition is more evenly distributed across auctions (the "demand distribution" effect). The latter effect exists because the low search costs facing online bidders make it relatively costless for them to shift their demand from auctions in which bidding competition is high to those in which bidding competition is expected to be low. Given that the positive price effects of demand expansion are relatively straightforward, we focus instead on the price effects of demand distribution by modeling the circumstances under which a more even distribution of bidder participation across auctions yields higher expected prices. We show that a more even distribution of bidder participation yields higher prices under fairly general conditions. However, this depends on the variability of bidder valuations for the objects being auctioned, the degree to which sellers can (and do) predict this a priori, and the average number of bidders per auction.

From a managerial perspective, the results have implications for sellers who use multiple auctions to sell products and who are interested in achieving higher prices. An obvious strategy for the seller is to attract more bidders to her auctions. It is nonobvious, however, that the seller can receive higher prices *without* increasing the number of bidders. Instead, she can often receive higher prices by distributing existing bidders more evenly across auctions. Possible strategies for doing this include spreading auctions across time or space and/or using promotional strategies to shape bidder participation. Testing which of these strategies is most effective is beyond the scope of our analysis but a fruitful avenue for future research. Engendering a more even distribution of bidder participation is a particularly good strategy if the seller cannot (or does not) forecast the variability of bidder valuations for each object a priori, which is likely to be the case for sellers whose inventory changes frequently or who sell objects for which buyer preferences change frequently. In addition to the relevance to sellers, our findings should be of interest to auction houses whose fees are based on a percentage of the selling price.

We tested the insights of our model using a unique data set of over 65,000 sequential auctions comprising over 10,000,000 vehicles in the wholesale used vehicle market. These auctions are conducted at physical market facilities, and bidders can submit bids physically or electronically (through an Internet application that streams live audio/video of the auctions). We show that physical bidders are more likely to bid early in the sequence, whereas electronic bidders are more likely to bid later in the sequence. We attribute this to electronic bidders having lower search costs than physical bidders. This makes it easier for electronic bidders to wait until later in the sequence to bid, when bidding competition from physical bidders lessens. The attrition among physical bidders and the converse for electronic bidders create a more even distribution of bidder participation across the auctions in the sequence. Consistent with the analytical model, we show that a more even distribution of bidder participation is associated with higher prices, but that this is moderated by the average number of bidders who bid on the vehicles in the sequence and whether the seller sequences the vehicles based on the variability of bidder valuations for them. We estimate that more even demand distribution (specifically, a reduction of one standard deviation in the standard deviation of bidders across the auctions in the sequence) yields a more than $10,000 increase (on average) in the total expected revenue from the auctions in the sequence.

5.2. Limitations

Because of space and scope considerations, we focus our analysis on how the distribution of bidder participation influences prices. We do not explore how other auction parameters, such as optimal reserve prices or changes to auction design, influence prices. Also, we do not model individual bidders' decisions to bid on a given vehicle, nor do we model bidders' decisions to bid physically or electronically. Instead, we observe the auctions in which bidders participate after having chosen a channel, and we use this information to measure the distribution of bidder participation across auctions. Future research could investigate bidders' decisions at a microlevel; some studies of this type have already been conducted (e.g., Bapna et al. 2010). Also, we cannot fully eliminate the possibility that relationships in our regressions are biased because of omitted variables or other sources of endogeneity, despite our use of multiple control variables and robustness checks. Similarly, if unobserved factors influence how sellers in our empirical application sequence vehicles, then this could confound our analysis in §§4.2.1 and 4.2.2, although we have controlled for the main sequencing methods identified by Grether and Plott (2009). Also, we use the regressions in §4.2.2 to examine the relationship between run number and the number of bidders after controlling for other variables—and to decompose this overall relationship to see how it differs based on bidder type (i.e., physical or electronic). We do not estimate a simultaneous equations system for the number of physical and electronic bidders, which is an opportunity for future research. Also, we used a proxy measure for the variability of bidder valuations per object. We believe this to be unavoidable given the archival nature of the data used in the empirical portion of our paper; measuring these variables with perfect precision would require an experimental approach. Data limitations also required us to impute the number of physical bidders. Potential measurement error in this variable represents a limitation of the empirical portion of this research, which we explore in the appendix.

Acknowledgments

This research has benefited from valuable comments from Lorin Hitt (department editor), the associate editor, the referees, and seminar participants at the Georgia Institute of Technology, the University of Georgia, the 2010 INFORMS Annual Meeting, the 2010 Workshop on Information Systems and Economics, and the 2011 INFORMS Conference on Information Systems and Technology. The authors thank the Georgia Institute of Technology and Purdue University for financial support.

Appendix

A.1. Constructing the Bid Log Variables

We extracted $Bids_{ij}$, $Bidders_{ij}$, $PBids_{ij}$, $PBidders_{ij}$, $EBids_{ij}$, $EBidders_{ij}$, $HighBid_{ij}$, and $StartingBid_{ij}$ from the bid log. Some institutional detail is necessary to understand the data contained within the bid log. If a bid is placed by a physical bidder, then a clerk at the auction facility enters the bid amount but not the identity of the physical bidder in the bid log. If a bid is placed by an electronic bidder, then the bid amount is automatically entered in the bid log, along with the ID of the bidder based on the credentials used to log into the electronic webcast application. The bidder ID is recorded for a winning bid, for both physical and electronic bidders. Thus, each row of a vehicle's bid log contains the bid amount, whether the bid was placed by a physical or electronic bidder, and the bidder ID for all electronic bids and for winning physical bids. Table A.1 illustrates.

For each vehicle j in sales event i, we extracted the starting bid ($StartingBid_{ij}$), high bid ($HighBid_{ij}$), and the sequence of physical and electronic bids ($BidPattern_{ij}$). For the bid log in Table A.1, $BidPattern_{ij} =$ "PEEPP." We constructed $Bidders_{ij}$ as the sum of the number of physical bidders ($PBidders_{ij}$) and electronic bidders ($EBidders_{ij}$). We counted the number of electronic bidders from the bid log. However, we could not count the number of physical bidders because

Table A.1 Example of the Bid Log for a Vehicle

Bid amount	Bid type	Bidder ID
10,000	P	
10,100	E	111
10,200	E	222
10,300	P	
10,400 (winning bid)	P	333

Note. P, physical; E, electronic.

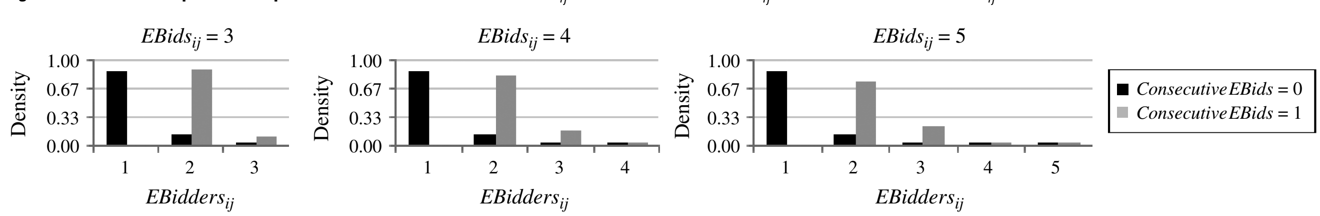
their IDs are not recorded (unless they place a winning bid). Thus, we imputed the number of physical bidders. Imputation was trivial for three cases that cover 17% of the sample. If a vehicle i had no physical bids (n =290,559), we set $PBidders_{ij} = 0$. If a vehicle had one physical bid ($n = 824{,}315$), we set $PBidders_{ij} = 1$. If a vehicle had two physical bids that were placed consecutively ($n = 678{,}620$), we set $PBidders_{ij} = 2$, because a bidder will not immediately outbid himself. For the remaining cases, we imputed the number of physical bidders for a vehicle based on the number and pattern of physical bids for that vehicle. We did this by assuming that the *unobserved* relationship between physical bids and bidders is analogous to the *observed* relationship between electronic bids and bidders (we justify this assumption below).

To implement this procedure, we leveraged not only the number of bids but also whether bids of the same type were placed consecutively. We created $ConsecutiveEBids_{ij}$, which is an indicator variable set to 1 for vehicles in which two electronic bids were placed consecutively. We created $ConsecutivePBids_{ij}$ analogously. We calculated the empirical distribution of $EBidders_{ij}$ for each combination of $EBids_{ij}$ and $ConsecutiveEBids_{ij}$. Figure A.1 shows the histograms for some of these distributions. As shown in the first panel of Figure A.1, if $EBids_{ij} = 3$ and $ConsecutiveEBids_{ij} = 1$, then $EBidders_{ij} = 2$ approximately 90% of the time. We used the distribution of $EBidders_{ij}$ conditional on $EBids_{ij}$ and $ConsecutiveEBids_{ij}$ as a proxy for the distribution of $PBidders_{ij}$ conditional on $PBids_{ij}$ and $ConsecutivePBids_{ij}$. We imputed $PBidders_{ij}$ by taking a draw from the appropriate distribution based on the values of $PBids_{ij}$ and $ConsecutivePBids_{ij}$. For example, if $PBids_{ij} = 3$ and $ConsecutivePBids_{ij} = 1$, then we imputed $PBidders_{ij} = 2$ approximately 90% of the time.

Next, we justify—both theoretically and empirically—the assumption that the mapping between electronic bids and bidders is a valid proxy for the mapping between physical bids and bidders. Theoretically, we first distinguish between two distinct bidder behaviors relevant to our context: (a) how a bidder chooses the auctions in which to participate, and (b) his bidding strategy once he has chosen an auction. As shown in the main text, whether a bidder is using the physical or electronic channel influences the auctions in which he chooses to participate. However, we do not believe that which channel he uses influences his bidding strategy once he has chosen an auction. We reason that a bidder's strategy will be to remain in the bidding until he either wins or reaches his valuation, regardless of which channel he is using. This is a rational strategy because he can achieve positive surplus by purchasing at or below his valuation. Furthermore, 70% of the sold vehicles in the sample were purchased by "multichannel" bidders, i.e., those who purchased vehicles via both channels. It seems unlikely that these multichannel bidders would dramatically change their bidding strategies when shifting between the two channels. Given similar bidding strategies across the channels, the mapping between electronic bids and bidders should be a good proxy for the mapping between physical bids and bidders.

Empirically, we validated this mapping by engaging in primary data collection at the physical auction facilities. One of the authors traveled to four different physical auction sites in the Southeast United States in January, March, and May of 2013 to observe the auctions and record the number of physical bidders for each vehicle. This required close cooperation with the auctioneers, because bidders' bid gestures are often quite subtle and targeted directly at the auctioneer. At the end of each auction, the auctioneer would indicate the number of physical bidders to the author. We then extracted the number of physical bids and the bid pattern from the bid logs for each auction we observed. This yielded a sample of 429 vehicles for which we had data on the number of physical bidders ($PBidders_{ij}$) and physical bids ($PBids_{ij}$). We dropped 35 observations due to coding/communication problems (e.g., the number of physical bidders exceeded the number of physical bids). Another 99 observations were for the three "trivial" cases discussed above (i.e., 0 physical bids, 1 physical bid, or 2 physical bids placed consecutively). We refer to the remaining 295 observations as the *supplementary sample*. Using the supplementary sample, we constructed empirical distributions for $PBidders_{ij}$ for each combination of $PBids_{ij}$ and $ConsecutivePBids_{ij}$. Because $ConsecutivePBids_{ij} = 1$ for 98% of vehicles in the full sample that had 2 or more physical bids, we focused on those distributions. We used Pearson's χ^2 test and Fisher's exact test to assess whether the empirical distributions from the supplementary sample matched the corresponding distributions used in the imputation procedure. Fisher's exact test may be the more appropriate of the

Figure A.1 Examples of Empirical Distributions of $EBidders_{ij}$ **Conditional on** $EBids_{ij}$ **and** $ConsecutiveEBids_{ij}$



two because the distributions often have low mass where the number of bidders is large relative to the number of bids.[23] The null hypothesis for both tests is that the distributions do not differ. As shown in Table A.2, we cannot reject the null hypothesis in most cases. Because the statistical power of these tests may be low given the size of the supplementary sample, we also show in Table A.2 that the percentage of times each number of bidders is observed closely matches the predicted percentage from the imputation procedure. Our results support the validity of our imputation procedure.[24]

Necessarily, imputation introduces some mismeasurement into $PBidders_{ij}$; sometimes we will overestimate it, and sometimes we will underestimate it. This does not pose any particular problem in the vehicle-level regression for $PBidders_{ij}$ (or $Bidders_{ij}$) reported in Table 2 because measurement error in the dependent variable is captured in the regression's error term (Hausman 2001, p. 59). A more serious concern is the possibility that measurement error in $PBidders_{ij}$ will lead to measurement error in $StDevBidders_i$ and $AvgBidders_i$ in the sales event regression. This is because measurement error in an explanatory variable can bias the coefficients. It should be noted that $StDevBidders_i$ and $AvgBidders_i$ are computed after aggregating $Bidders_{ij}$ across vehicles in sales event i. Thus, even if $Bidders_{ij}$ is mismeasured, $StDevBidders_i$ and $AvgBidders_i$ may not be, because the measurement error in $Bidders_{ij}$ might "wash out" when aggregated (e.g., see Cameron and Trivedi 2005, p. 899).[25] Despite this possibility, we investigated the effects of potential measurement error in the sales event regressions.

Our primary approach was to adjust the coefficients and standard errors from the sales event regressions by the variance/covariance matrix of the measurement errors. This permits consistent estimation (Fuller 1987). We estimated the variance/covariance matrix as follows. We randomly drew 158 vehicles from the supplementary sample described above (we used 158 because that is the mean of $VehiclesOffered_i$ in our empirical analysis). We refer to

[23] A common rule of thumb for Pearson's χ^2 test is that each discrete point in the distribution contain at least 5 observations, although DeGroot and Schervish (2002, p. 537) state that 1.5 is satisfactory.

[24] We considered whether our validation of the imputation procedure using data from 2013 was valid given that we use the imputation procedure on data from 2005 to 2010. We do not believe this to be a problem, because the institutional context is fairly stable over time: the basic auction procedure has remained intact and the same bidders participate year over year. We also checked this empirically. In addition to receiving the bid logs for the 429 vehicles for which we manually observed the number of physical bidders, we also received the bid logs for an additional 6,080 vehicles auctioned during our site visits that we didn't observe. We compared the empirical distributions of $EBidders_{ij}$ for each combination of $EBids_{ij}$ and $ConsecutiveEBids_{ij}$ from these data to the distributions we used for imputation using Pearson's χ^2 test and Fisher's exact test. We found no significant differences, indicating that the imputation distributions that we use are stable over time.

[25] For example, if the true number of bidders in each auction in a sales event is (2, 2, 3), then $AvgBidders_i = 2.33$ and $StDevBidders_i = 0.57$. If our imputation of the number of bidders yields any combination of two 2s and one 3 (i.e., (2, 2, 3), (2, 3, 2), or (3, 2, 2)), then our measures of $AvgBidders_i$ and $StDevBidders_i$ will be correct.

Table A.2 Comparison of Empirical Distributions for $PBidders_{ij}$ for Each Level of $PBids_{ij}$ from the Supplementary Sample to the Corresponding Distributions Used in the Imputation Procedure

# of bids = 3: # of bidders	Times observed	Predicted %
2	8 (89%)	90
3	1 (11%)	10
$\chi^2_{(1)} = 0.01$ ($p = 0.92$)		
Fisher's exact: $p = 1.00$		

# of bids = 4: # of bidders	Times observed	Predicted %
2	9 (64%)	81
3	4 (29%)	17
4	1 (7%)	2
$\chi^2_{(2)} = 6.31$ ($p = 0.04$)		
Fisher's exact: $p = 0.06$		

# of bids = 5: # of bidders	Times observed	Predicted %
2	12 (80%)	75
3	2 (13%)	22
4	1 (7%)	3
5	0 (0%)	0
$\chi^2_{(3)} = 1.42$ ($p = 0.70$)		
Fisher's exact: $p = 0.36$		

# of bids = 6: # of bidders	Times observed	Predicted %
2	14 (67%)	68
3	6 (29%)	26
4	1 (5%)	5
5	0 (0%)	1
6	0 (0%)	0
$\chi^2_{(4)} = 0.16$ ($p = 0.99$)		
Fisher's exact: $p = 0.93$		

# of bids = 7: # of bidders	Times observed	Predicted %
2	8 (62%)	63
3	4 (31%)	29
4	1 (8%)	7
5	0 (0%)	1
6	0 (0%)	0
7	0 (0%)	0
$\chi^2_{(5)} = 0.19$ ($p = 0.99$)		
Fisher's exact: $p = 0.92$		

# of bids	8	9	10	11	12	13	14	15	16	17+
Times observed	14	12	14	15	25	10	16	13	15	89
χ^2 (p-value)	3.7 (0.71)	5.6 (0.59)	11.2 (0.2)	1.8 (0.99)	6.8 (0.75)	1.0 (0.99)	2.9 (0.99)	1.8 (1.00)	1.9 (1.00)	n/a
Fisher's exact: p-value	0.25	0.11	0.02	0.79	0.20	0.93	0.70	0.75	0.87	n/a

Notes. The column labeled as "# of bids = 3" shows that for the vehicles that received 3 physical bids in the supplementary sample, we observed 8 to have two physical bidders and 1 to have three physical bidders. The "Predicted %" column shows the expected percentage of each outcome from the imputation procedure. The χ^2 statistic shows the results of Pearson's χ^2 test of whether the observed distribution differs from the distribution used for imputation. The Fisher's exact statistic is analogous. Similar tables are provided for 4, 5, 6, and 7 physical bids. For 8+ bids, we present only the χ^2 and Fisher statistics to conserve space. The percentages shown may not sum to 100% due to rounding.

Table A.3 Results of the Sales Event Regression Aafter Adjusting for Measurement Error

	Dep. var.: $AvgHighBid_j$	Dep. var.: $AvgHighBid_j$
	Coef. (robust SE)	Coef. (robust SE)
β_1: $AvgValuation_i$	0.961 (0.001)***	0.961 (0.001)***
β_2: $AvgMileage_i$	0.003 (0.002)	0.003 (0.002)
β_3: $AvgCondition_i$	1,210.408 (10.636)***	1,222.650 (10.545)***
β_4: $VehiclesOffered_i$	0.950 (0.048)***	0.948 (0.048)***
β_5: Day_i	−0.107 (0.008)***	−0.111 (0.008)***
β_6: $StDevBidders_i$	−929.655 (34.466)***	−732.758 (24.937)***
β_7: $AvgBidders_i$	649.832 (15.698)***	598.466 (14.382)***
β_8: $ValuationVar\_SalesEvent_i$	0.009 (0.003)***	0.008 (0.003)***
β_9: $SequenceByValDist_i$	16.723 (10.393)	15.087 (10.380)
β_{10}: $StDevBidders_i * AvgBidders_i$		69.589 (9.929)***
β_{11}: $StDevBidders_i * ValuationStDev\_SalesEvent_i$		−0.020 (0.015)
β_{12}: $StDevBidders_i * SequenceByValDist_i$		274.518 (51.007)***
β_0: *Intercept*	−3,657.72 (39.953)***	−3,677.439 (40.097)***
n	65,178	65,178
R^2	n/a	n/a
Wald F-statistic (d.f.)	138,291 (9, 65,708)	108,551 (12, 65,705)

***$p < 0.10$.

these 158 vehicles as the *supplementary sales event*. Because we observe the true value of $PBidders_{ij}$ for each vehicle in the supplementary sales event, we can compute the true values of $StDevBidders_i$, $AvgBidders_i$, and the $StDevBidders_i$ interaction terms. We also imputed $PBidders_{ij}$ for each vehicle in the supplementary sales event and used those values to compute the imputed values for $StDevBidders_i$, $AvgBidders_i$, and the $StDevBidders_i$ interaction terms.[26] We subtracted the imputed values from the true values to get the measurement errors. We repeated this procedure to create 10,000 "bootstrapped" supplementary sales events and 10,000 instances of each measurement error. We computed the variance/covariance matrix of the measurement errors using these 10,000 instances of each error. We reestimated the sales event regressions (without the seller fixed effects)[27] using this matrix to adjust for the measurement errors. Results appear in Table A.3. The coefficients for $StDevBidders_i$ and $AvgBidders_i$ are larger in absolute value than those reported in Table 4, suggesting coefficient attenuation. (They are also larger if we reestimate the Table 4 results without the seller fixed effects.) This is not particularly surprising, because measurement error is known to cause attenuation, although the direction of the bias is equivocal a priori when more than one explanatory variable is measured with error. This suggests that our results in Table 4 may be conservative, such that the economic significance of our results may be even greater than what we report in the main text.

A potential limitation of the above analysis is that the variance/covariance matrix of the measurement errors recovered from the supplementary sample may differ from that of the main sample. For robustness against this possibility, we implemented an alternative approach to account for the potential measurement error by using instrumental variables. Wooldridge (2002, Chap. 4.4.2, pp. 105–106) referred to this approach as the *multiple indicator solution*. In this approach, we considered both $AvgBidders_i$ and $AvgBids_i$ to be observed indicators of the unobserved true value of $AvgBidders_i$ (with $StDevBidders_i$ and $StDevBids_i$ analogous); i.e., $AvgBidders_i = AvgBidders_i^* + \alpha_1$ and $AvgBids_i = AvgBidders_i^* + \alpha_2$, where $AvgBidders_i^*$ is the unobserved true value and α_1 and α_2 are the measurement errors. $AvgBids_i$ ($StDevBids_i$) can be used as an instrument for $AvgBidders_i$ ($StDevBidders_i$) in the sales event regressions if (a) the measurement errors α_1 and α_2 are uncorrelated with $AvgBidders_i^*$ and the other explanatory variables, (b) $AvgBids_i$ is correlated with $AvgBidders_i$, (c) $AvgBids_i$ has no influence on $AvgHighBid_i$ except through its relationship with $AvgBidders_i^*$, and (d) α_1 and α_2 are uncorrelated with each other. The assumptions in (a) are the standard ones for errors-in-variables models, they are almost always maintained in practice, and they seem reasonable for our context. (b) is satisfied; the correlation between $AvgBids_i$ and $AvgBidders_i$($StDevBids_i$ and $StDevBidders_i$) is 0.75 (0.66). (c) seems reasonable, given that the theoretical changes to $AvgHighBid_i$ modeled in the regression pertain to bidders, not bids. (d) seems reasonable because the correlation between $AvgBids_i$ and $AvgBidders_i$ stems from their common correlation with $AvgBidders_i^*$, not from correlation between α_1 and α_2. The instrumental variable results (available from the authors) are consistent with the adjusted results reported in Table A.3; they also show coefficient attenuation.

A.2. Proof of Concavity of Prices

Lemma A1. *$Price_m(x)$ is strictly concave in x.*

Proof. Recall that $Price_m(x)$ is the expectation of the $(x-1)$st-order statistic from the set of bidder valuations

[26] To create the true and imputed $StDevBidders_i * ValuationStDev\_SalesEvent_i$ and $StDevBidders_i * SequenceByValDist_i$ interaction terms, we randomly drew a value for $ValuationStDev\_SalesEvent_i$ and $SequenceByValDist_i$ from the main sample of 65,718 sales events.

[27] Including the seller fixed effects creates computational problems when attempting to invert the matrices needed for estimation of the coefficients and their standard errors (for the formulas, see Fuller 1987, §3.1.3, pp. 199–202).

Table A.4 Summary Statistics for How Sellers Sequence Vehicles in Sales Events

		Proportion of sales events in which both patterns occur[a]											
Pattern	Mean	1	2	3	4	5	6	7	8	9	10	11	12
1. $SequenceByValuation(\downarrow)_i$	0.44	—											
2. $SequenceByValuation(\uparrow)_i$	0.15	—	—										
3. $SequenceByValuation_i = 0$	0.41	—	—	—									
4. $SequenceByMileage(\downarrow)_i$	0.19	0.04	0.08	0.06	—								
5. $SequenceByMileage(\uparrow)_i$	0.46	0.34	0.03	0.10	—	—							
6. $SequenceByMileage_i = 0$	0.35	0.06	0.04	0.25	—	—	—						
7. $SequenceByCondition(\downarrow)_i$	0.35	0.25	0.03	0.07	0.04	0.25	0.06	—					
8. $SequenceByCondition(\uparrow)_i$	0.13	0.04	0.05	0.04	0.06	0.04	0.04	—	—				
9. $SequenceByCondition_i = 0$	0.53	0.16	0.08	0.29	0.09	0.18	0.26	—	—	—			
10. $SequenceBySupply(\downarrow)_i$	0.27	0.16	0.04	0.07	0.05	0.16	0.06	0.12	0.03	0.11	—		
11. $SequenceBySupply(\uparrow)_i$	0.22	0.09	0.06	0.07	0.07	0.09	0.06	0.07	0.05	0.10	—	—	
12. $SequenceBySupply_i = 0$	0.51	0.19	0.05	0.27	0.06	0.21	0.24	0.16	0.05	0.31	—	—	—

Notes. The number of sales events is 65,718. The standard deviation, median, min, and max are withheld for economy of presentation. Because each variable is binary, the min and max are always 0 and 1; the median is 1 if the mean > 0.5, 0.5 if the mean = 0.5, and 0 otherwise; and the standard deviation is $\sqrt{(n^*\mu^*(1-\mu))/(n-1)}$, where μ is the mean and $n = 65{,}718$.

[a]These figures represent the proportion of sales events in which both patterns are evident. For example, in 25% of the sales events, there was a negative relationship between $RunNumber_{ij}$ and $Valuation_{ij}$ (pattern 1) and a negative relationship between $RunNumber_{ij}$ and $Condition_{ij}$ (pattern 7).

for object m when there are x bidders, i.e., when there are x draws from the $F_m(\cdot)$ distribution. We denote this as $Price_m(x) = \mu_{x-1:\,x}$. To prove concavity, we need the following:

$$\begin{aligned} \mu_{x-1:x} - \mu_{x:x+1} &> \mu_{x-2:x-1} - \mu_{x-1:x}, \\ 2\mu_{x-1:x} &> \mu_{x-2:x-1} + \mu_{x:x+1}. \end{aligned} \tag{4}$$

We will use the following order statistics relation (David and Nagaraja 2003, p. 45):

$$\mu_{r:x} = \sum_{i=r}^{x} \frac{(-1)^{i-r}(i-1)!}{(r-1)!(i-r)!} \frac{x!}{i!(x-i)!} \mu_{i:i}.$$

After applying that relation to Equation (4), the condition for concavity is as follows:

$$\begin{aligned} &x(-\mu_{x-2:x-2} + 3\mu_{x-1:x-1} - 3\mu_{x:x} + \mu_{x+1:x+1}) \\ &\quad > (-\mu_{x-2:x-2} + 2\mu_{x-1:x-1} - \mu_{x:x}). \end{aligned} \tag{5}$$

Recall that the distribution of bidder valuations has an increasing hazard rate. We know from Barlow and Proschan (1966) that the normalized spacing for distributions with increasing hazard rates have decreasing stochastic orders (also, see Kochar and Kirmani 1995, p. 48). This, in turn, implies that the expectations for the normalized spacing satisfy the following:

$$\begin{aligned} D_i^* &= (\eta - i + 1)(\mu_{i:\eta} - \mu_{i-1:\eta}) \\ &\geq D_{i+1}^* = (\eta - i)(\mu_{i+1:\eta} - \mu_{i:\eta}) \end{aligned} \tag{6}$$

for $i = 1, \ldots, \eta - 1$.

By setting $i = x - 1$ and $\eta = x + 1$ in Equation (6) and using that relation in Equation (5), we get

$$\begin{aligned} &x(-\mu_{x-2:x-2} + 3\mu_{x-1:x-1} - 3\mu_{x:x} + \mu_{x+1:x+1}) \\ &\quad > (-\mu_{x-2:x-2} + 2\mu_{x-1:x-1} - \mu_{x:x}) \\ &\qquad + (-\mu_{x-1:x-1} + 2\mu_{x:x} - \mu_{x+1:x+1}). \end{aligned} \tag{7}$$

Note that Huang (1998) shows that the summation of terms within each set of parentheses is positive. This allows us to remove the second parenthetical term from Equation (7). This yields Equation (5) and completes the proof.

A.3. Seller Sequencing Strategies

Table A.4 summarizes the *SequenceBy*… variables discussed in §4.2.1. Table A.4 shows substantial heterogeneity in how sellers sequence vehicles, which is consistent with Grether and Plott (2009). For example, sellers sequenced vehicles from high to low $Valuation_{ij}$ in 44% of the sales events and from low to high in 15%. In the other 41% of the sales events, they appear to have ignored $Valuation_{ij}$. In this case, they sometimes sequenced vehicles by $Mileage_{ij}$ instead: 16% of sales events have $SequenceByValuation_i = 0$ and either $SequenceByMileage(\downarrow)_i = 1$ or $SequenceByMileage(\uparrow)_i = 1$ (see the 4th and 5th rows, 3rd column in Table A.4). There are 9,542 sales events (14.5%) in which $RunNumber_{ij}$ is not related to the other variables, i.e., in which there is no discernible pattern.

References

Allison P (2012) When can you safely ignore multicollinearity? *Statistical Horizons* (blog) (September 10), http://www.statisticalhorizons.com/multicollinearity.

Arnold BC (1987) *Majorization and the Lorenz Order: A Brief Introduction* (Springer, Berlin).

Arora A, Greenwald A, Kannan K, Krishnan R (2007) Effects of information revelation policies under market structure uncertainty. *Management Sci.* 53(8):1234–1248.

Ashenfelter O (1989) How auctions work for wine and art. *J. Econom. Perspect.* 3(3):23–36.

Bajari P, Hortacsu A (2003) The winner's curse, reserve prices and endogenous entry: Empirical insights from eBay auctions. *RAND J. Econom.* 34(2):329–355.

Bapna R, Dellarocas C, Rice S (2010) Vertically differentiated simultaneous vickrey auctions: Theory and experimental evidence. *Management Sci.* 56(7):1074–1092.

Bapna R, Chang SA, Goes P, Gupta A (2009) Overlapping online auctions: Empirical characterization of bidder strategies and auction prices. *MIS Quart.* 33(4):763–784.

Barlow RE, Proschan F (1966) Inequalities for linear combinations of order statistics from restricted families. *Ann. Math. Stat.* 37(6):1574–1592.

Baye MR, Morgan J, Scholten P (2006) Information, search, and price dispersion. Hendershott T, ed. *Handbook on Economics and Information Systems* (Elsevier, Amsterdam), 323–376.

Bernhardt D, Scoones D (1994) A note on sequential auctions. *Amer. Econom. Rev.* 84(3):653–657.

Brendstrup B, Paarsch HJ (2006) Identification and estimation in sequential, asymmetric, English auctions. *J. Econometrics* 134(1):69–94.

Brown JR, Goolsbee A (2002) Does the Internet make markets more competitive? Evidence from the life insurance industry. *J. Political Econom.* 110(3):481–507.

Brynjolfsson E, Smith MD (2000) Frictionless commerce? A comparison of Internet and conventional retailers. *Management Sci.* 46(4):563–585.

Cameron AC, Trivedi P (2005) *Microeconometrics Methods and Applications* (Cambridge University Press, New York).

Clemons EK, Hann IH, Hitt LM (2002) Price dispersion and differentiation in online travel: An empirical investigation. *Management Sci.* 48(4):534–549.

David HA, Nagaraja HN (2003) *Order Statistics*, 3rd ed. (John Wiley & Sons, Hoboken, NJ).

DeGroot MH, Schervish MJ (2002) *Probability and Statistics*, 3rd ed. (Addison Wesley, Boston).

Echambadi R, Hess JD (2007) Mean-centering does not alleviate collinearity problems in moderated multiple regression models. *Marketing Sci.* 26(3):438–445.

Edelman B, Ostrovsky M, Schwarz M (2007) Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *Amer. Econom. Rev.* 97(1):242–259.

Engelbrecht-Wiggans R (1994) Sequential auctions of stochastically equivalent objects. *Econom. Lett.* 44(1–2):87–90.

Fuller WA (1987) *Measurement Error Models* (John Wiley & Sons, New York).

Genesove D (1995) Search at wholesale auto auctions. *Quart. J. Econom.* 110(1):23–49.

Goes PB, Karuga GG, Tripathi AK (2010) Understanding willingness-to-pay formation of repeat bidders in sequential online auctions. *Inform. Systems Res.* 21(4):907–924.

Greenwald A, Kannan K, Krishnan R (2010) On evaluating information revelation policies in procurement auctions: A markov decision process approach. *Inform. Systems Res.* 21(1):15–36.

Grether DM, Plott CR (2009) Sequencing strategies in large, competitive, ascending price automobile auctions: An experimental examination. *J. Econom. Behav. Organ.* 71(2):75–88.

Hausman J (2001) Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *J. Econom. Perspect.* 15(4):57–67.

Huang JS (1998) Sequence of expectations of maximum-order statistics. *Statist. Probab. Lett.* 38(2):117–123.

Iacus SM, King G, Porro G (2011) Causal inference without balance checking: Coarsened exact matching. *Political Anal.* 20(1):1–24.

Imbens G (2004) Nonparametric estimation of average treatment effects under exogeneity: A review. *Rev. Econom. Statist.* 86(1):4–29.

Kannan K (2010) Declining prices in sequential auction with complete revelation of bids. *Econom. Lett.* 108(1):49–51.

Kannan K (2012) Effects of revelation policies under cost uncertainty. *Inform. Systems Res.* 23(1):75–92.

Kazumori E, McMillan J (2005) Selling online vs. live. *J. Indust. Econom.* 53(4):543–569.

Kochar SC, Kirmani SNUA (1995) Some results on normalized spacings from restricted families of distributions. *J. Statist. Planning Inference* 46(1):47–57.

Koppius O, van Heck E, Wolters M (2004) The importance of product representation online: Empirical results and implications for electronic markets. *Decision Support Systems* 38(2):161–169.

Kuruzovich J, Viswanathan S, Agarwal R (2010) Seller search and market outcomes in online auctions. *Management Sci.* 56(10):1702–1717.

Lee HG (1998) Do electronic marketplaces lower the price of goods? *Comm. ACM* 41(1):73–80.

Myerson RB (1981) Optimal auction design. *Math. Oper. Res.* 6(1): 58–73.

NAAA (National Auto Auction Association) (2010) NAAA 2010 annual review. Report, http://www.naaa.com/pdfs/2010_NAAAAnnualReview.pdf.

NAAA (National Auto Auction Association) (2012) NAAA 2012 annual review. Report, http://www.naaa.com/pdfs/2012AnnualReview.pdf.

Overby E, Jap S (2009) Electronic and physical market channels: A multiyear investigation in a market for products of uncertain quality. *Management Sci.* 55(6):940–957.

Smith GE, Venkatraman MP, Dholakia RR (1999) Diagnosing the search cost effect: Waiting time and the moderating impact of prior category knowledge. *J. Econom. Psych.* 20(3):285–314.

Stahl D (1989) Oligopolistic pricing with sequential consumer search. *Amer. Econom. Rev.* 79(4):700–712.

Varian HR (1980) A model of sales. *Amer. Econom. Rev.* 70(4): 651–659.

Wooldridge JM (2002) *Econometric Analysis of Cross Section and Panel Data* (MIT Press, Cambridge, MA).