EXPLAINING PRICE DISPERSION IN ONLINE AUCTIONS WITH SEARCH FRICTIONS[1]

by

Joseph Uri Podwol
Department of Economics
Cornell University

Henry S. Schneider
Johnson Graduate School of Management
Cornell University

June 19, 2012

Very preliminary and incomplete.

This version may contain significant mistakes, omissions, and unclear passages.

Abstract

A standard explanation for price dispersion in fixed-price markets is search frictions faced by buyers. We incorporate this insight into a model of a competing-auctions market to explain price dispersion in this alternate setting. In our model, the search costs of searching for competing auctions causes a subset of buyers to limit their search to only the most visible auctions. Price dispersion arises as these buyers bid up the price in the most visible auctions, above the prices for identical items in less visible auctions. Predictions of the model distinguish auction markets from those of fixed-price sales. We test predictions from the model using data on eBay auctions and find that competing auctions that are less likely to be identified by the same keyword search have more dispersed prices.

[1] We thank Nate Woody for excellent programming assistance, and the Cornell University Institute for Social Sciences for financial support.

1. INTRODUCTION

In his seminal 1980 article on price dispersion, Varian writes that "The law of one price is no law at all." Indeed, price dispersion has been documented in a range of markets from consumer products, such as automobiles and retail gasoline, to industrial inputs, such as ready-mix concrete and government purchases of anthracite coal.[2] A leading explanation in the literature is the presence of *unaware* consumers who, whether due to information acquisition costs or loyalty to specific sellers, are imperfectly informed about the full set of competing prices (e.g., Varian 1980, Rosenthal 1980, Narasimhan 1988, Stahl 1989). Price dispersion arises as firms randomize between low-volume, high-margin sales to buyers who are unaware of competing sellers ("unaware" buyers), and high-volume, low-margin sales to buyers who are aware of competing sellers ("aware" buyers).[3]

The existing literature focuses on markets with fixed-price sellers and is not as well suited to understanding price dispersion in markets with competing auctions, such as eBay. Using data on eBay auctions for new movie-DVD auctions with standard shipping (described later), we group auctions by movie title, format (e.g., Blu-ray), and transaction date. Table 1 reports a measure of price dispersion, which is the standard deviation of auction ending prices as a fraction of the mean daily price ranges. Considerable price dispersion is apparent, with the measure ranging from 5 percent to 21 percent.[4] In the current study, we seek to explain this price

[2] See Table 1 in Baye, Morgan and Scholten (2006) for a more complete list.
[3] This type of price dispersion is not common to all models that incorporate information acquisition costs. Salop and Stiglitz (1977), Carlson and McAfee (1983), MacMinn (1980), Reinganum (1979) and Spulber (1995) model price dispersion in pure strategies. These models, however, are unable to explain why price dispersion would persist over time since consumers might reasonably be expected to eventually identify low-price firms. Baye and Morgan (2001) and Burdett and Judd (1983) show that consumers needn't be heterogeneous for price dispersion to result, however, heterogeneity in search costs seems to be the more plausible modeling assumption.
[4] The mean coefficient of variation across all title-editions (the bottom row of Table 1) is 0.12. This result is similar to that of Einav, Kuchler, Levin and Sundaresen (2011), who find a mean coefficient of variation of 0.11 across all non-auto or real estate categories on eBay. Thus, it does not appear that DVDs are at all unusual in the level of price dispersion observed.

dispersion by incorporating insights from the existing fixed-price literature into a model of competing auctions.

The distinction between fixed-price and auction markets is compelling due to the basic differences in their respective price-setting mechanisms. In auctions markets, prices are determined through a bidding process that is distinct from the fixed-price mechanisms mentioned above. For example, auction prices may directly reflect idiosyncratic buyer valuations, which could increase dispersion relative to fixed prices. Alternatively, auction prices may be determined by buyers' cross-bidding across auctions in incremental bid amounts, which may reduce price dispersion (Peters and Severinov 2006). This suggests an enriched model of the competing-auctions can shed light on price setting in this type of market.

Our analysis builds on the theoretical work of Peters and Severinov (2006) who model a market of simultaneous ascending-price auctions for identical items. In their model, all buyers are aware of all competing auctions and may bid multiple times across all of the auctions. Peters and Severinov (2006) demonstrate the existence of an equilibrium in which bidders bid the lowest allowable bid in the auction with the lowest standing price, which ensures that all auctions that result in sale have the same ending price.[5] The uniform-price equilibrium of Peters and Severinov (2006) establishes the auction-analogue of the law of one price.

We use the model in Peters and Severinov (2006) as a baseline to which we add search frictions as in Varian (1980). Bidders in our model incur search costs to identify the existence and prices of all competing auctions, which in practice involves bidders sorting through auction listings with similar titles and reading their detailed product specifications. Since popular items

[5] Huang et al. (2008) and Hendricks, Onur and Wiseman (2012) consider models where auctions overlap for some of their durations, but end at different times. This may be a more realistic representation of the eBay environment, but the focus of these models is not on cross-bidding, which is central for examining price dispersion.

on eBay can have many hundreds of competing listings, this process may involve significant time and effort.

One way in which bidders can reduce search costs is to include additional details about the desired item in their search string. By default, eBay's search algorithm generally requires every word in the bidder's search string to appear in the seller's listing title for the listing to appear in the bidder's search results. It follows that buyers who use more restrictive search criteria will be exposed to fewer listings, thereby becoming unaware of auctions not captured by their search.[6] Price dispersion arises as unaware buyers bid up the price in auctions they are aware of irrespective of the prices in competing auctions they are unaware of.[7]

The distinction between auction and fixed-price markets is highlighted by the disparate predictions regarding the fraction of buyers who are aware. The top panel of Figure 1 illustrates our model's prediction that price dispersion decreases monotonically with the fraction of buyers who are aware. As the fraction of aware buyers decreases, the unaware buyers have a larger affect on the prices of individual auctions and, in expectation, prices deviate more from the uniform-price equilibrium of Peters and Severinov (2006). In contrast, the bottom panel of Figure 1 shows the inverse-U relationship between price dispersion and the fraction of aware buyers as predicted by Varian (1980), Stahl (1989), and others. In these fixed-price settings, as

[6] Another approach to narrowing the search results is to use a set of navigation check boxes on the left side of the eBay screen. For example, a user looking for a new version of *Casino Royale* could type "Casino Royale" into the search bar; then click the left-side option "new" condition to exclude used items. This approach may still require the buyer to inspect several listings to distinguish between the two-disc full-screen version, two-disc widescreen version, the three-disc collector's edition, the 2002 DVD release of the 1967 movie called "Casino Royale" and 2008 DVD release of the 40th anniversary edition of the 1967 movie. The user must also distinguish between the plethora of seller characteristics such as reputation score, auction characteristics such shipping fee, and so on. Since this approach is still very time intensive, our analysis does not distinguish between this approach and manually inspecting the full set of listings. In contrast, typing the words "new" and "2008" immediately reduce the number of listings to inspect and hence reduce search costs.

[7] In practice, there are other frictions besides the one identified that can lead to price dispersion. These are discussed in Section 5.

the fraction of aware buyers approaches zero, sellers extract increasingly-more surplus. In the limit, all sellers charge the monopoly price, and price dispersion is zero.

In the second part of the paper, we test the theoretical predictions of the model using data on eBay auctions for movie DVDs. In our primary analysis, we test the prediction that price dispersion increases in the fraction of bidders who are unaware. We do not observe bidders' search decisions and hence we cannot distinguish between aware and unaware bidders directly. However, bidders will have included certain key descriptors in their search strings, and we predict that a fraction of bidders will be aware of auctions with these descriptors in their listing titles but unaware of auctions that do not. For example, bidders looking for a new version of a particular movie may specify this in their search string. Bidders may also indicate their preferred format by including the word "dvd" or "blu-ray" or "blu-ray disc." In this way, listing wording differences become a proxy for the fraction of bidders who are unaware.[8] Consistent with the theory, we find that among pairs of competing auctions, those whose listing titles differ in their inclusion of "new" or "disc" exhibit larger differences in ending prices.[9]

We further analyze pairs of auctions that differ as to the inclusion of one of the terms "new," "disc" or the auction's format ("DVD" or "Blu-ray"), which provides additional support for the theory. Consistent with our treatment of listing-wording differences as a proxy for unaware buyers, we find that auctions that include these descriptors in their listing titles: 1) Have higher transaction prices; and 2) Are more likely to receive a bid from buyer when the auction has the higher standing price of the two. To the extent that the inclusion of these descriptors in the listing title increases the number of buyers that are aware of the auction, the theory predicts

[8] Podwol and Schneider (2012) analyze the effects of wording differences on price comparisons between auction prices and posted ("Buy-It-Now") prices. We find that the auction price is more likely to exceed the Buy-It-Now price when the auction listing title includes words not in the Buy-It-Now listing title.
[9] To address the potential endogeneity of listing wording, we use an instrumental variables approach.

that these auctions should also have higher starting prices. We find that auctions that include the words "new" and "disc" in the listing title do in fact have higher starting prices on average.

To our knowledge, this is the first study to explicitly model price dispersion in the auction setting. Since the price-setting mechanism in auction markets is distinct from fixed-price markets, the existing models of price dispersion do not apply directly. Another paper that models the eBay environment is Hendricks, Onur, and Wiseman (2012). This paper investigates the eBay environment as sequential auctions in order to explain the effect of late-arriving bidders on price, but this framework is not as well suited to understanding price dispersion. Competition between simultaneous auctions is modeled in Wolinsky (1988), McAfee (1993), and others, but bidders are restricted to bid in one auction only, which again is not as well suited to understanding price dispersion in the current context. A contribution of the current paper is to incorporates key insights from the literature on price dispersion in fixed-price markets into the model of competing but frictionless auctions market in Peters and Severinov (2006). We provide a range of empirical results in support of the theoretical findings.

Our results also build on several empirical studies of search frictions in online auctions. Haruvy and Popkowski (2010) show that price dispersion decreases when bidders are given explicit incentives to monitor competing auctions.[10] Our study offers evidence about why bidders are unaware in the first place: search frictions due to factors such as listing-title wording differences. Anwar, McMillan and Zheng (2006) and Haruvy and Popkowski (2010) show that auctions won by buyers who bid in multiple competing auctions ("cross-bidders") have lower prices. The lower prices are a prediction of aware bidding that arises directly from our model.

[10] The search incentive was created by announcing on the auction listing page that the seller (i.e., the authors) would waive the shipping charge if the final price in that auction ended up lower than in any auction ending within 30 minutes of the auction in question (which would have included the matched pair).

The paper proceeds as follows. Section 2 describes the data. Section 3 provides basic facts about how listing wording varies by seller, which motivates the theoretical model. Section 4 provides the model and empirical predictions about price dispersion. Section 5 contains the empirical tests of the predictions. Section 6 concludes.

2. DATA

We collected data from eBay using a Java query tool, which we used to search the title and body of listings for certain movie titles. In selecting movie titles for our study, we began with the 25 best-selling DVDs according to Billboard magazine for August and September of 2008, then eliminated all non-movie DVDs (namely television series) leaving us with 17 movie titles. For each of these movie titles, we collected data on all auction and fixed-price Buy-It-Now (BIN) listings that were active between September and November 2008 for standard-format DVDs as well as Blu-ray discs. We designed our search to be as inclusive as possible to capture all listings we expected at least some eBay users to identify in their searches. Our listing search procedures are described in the Appendix.

Our primary dataset consists of bid-level data on completed auctions. A supplementary data set consists of all auctions and BIN listings regardless of whether a transaction occurred and is used to construct histories for each of the sellers in our sample. For the primary dataset, we record item characteristics such as title, condition (new and used), and whether the DVD is a special edition or widescreen version;[11] seller characteristics such as feedback score, whether

[11] During the sample period, eBay expanded the condition variable from "New" and "Used" to five categories from "Brand New" to "Acceptable." We reclassify the categories to "New" and "Used" and drop listings not reporting condition (0.05 percent of listings).

they carry a Powerseller designation and whether they have an eBay store;[12] shipping charge and type (e.g., priority); auction characteristics such as the entire listing title, the starting price, start and end time, and whether or not a secret reserve was used; and bid characteristics, including amount and time of each bid, whether the bid is an automatic proxy bid or actual bid;[13] bidder characteristics, notably bidder feedback score. After each transaction, the buyer can evaluate the seller, and vice versa, with a positive (+1), negative (-1), or neutral feedback (0); the feedback score is the sum of these feedbacks. In line with previous work, we use these feedback scores as measures of buyer and seller experience.

Sellers are identified uniquely by their seller-id, which we use to track the wording of the seller's other listings which were captured by our Java query tool. Upon dropping listings offering multiple items, we construct sellers' listing histories by grouping 4,385 auction and 2,253 BIN listings by seller-id. From there, we further dropped auctions that did not result in a sale, were privately listed or did not otherwise allow us to identify buyers; those with a shipping charge greater than $10 as they reflect significant outliers; and those for which the shipping charge was not provided. This leaves us with our primary data set of 15,036 bids (excluding proxy bids) from 1,380 distinct bidders, within 2,736 completed auctions conducted by 1,504 distinct sellers.

3. STYLIZED FACTS ON THE WORDING OF LISTING TITLES

[12] Inclusion in eBay's Powerseller program requires participating sellers to meet certain volume and quality requirements for which they receive additional services from eBay. An eBay store allows a seller to showcase items without having them listed as an auction or BIN listing.

[13] The proxy bidding system works as follows: when a bid is placed by someone other than the high bidder, the high bidder's bid is automatically increased up to his actual bid amount if the new bid exceeds his bid, or up to the new bid plus the minimum increment if the new bid is less than his actual bid. A bid placed by the high bidder has no effect on standing price, but increases the maximum amount that the proxy bidding system will bid for him. We distinguish between *proxy bids*, which are bids automatically placed up to a buyer's maximum, and *actual bids*, which are the amounts actually entered by the bidder, which serve as the maximum proxy bid.

We begin by providing an overview of seller behavior with respect to listing wording. The takeaways of this section are: 1) The inclusion of certain key descriptors is an important component of the seller's strategy set; 2) Between-seller differences explain much of the observed differences between listings as to the inclusion of key descriptors; and 3) The length of the movie title exacerbates between-seller differences by requiring otherwise-similar sellers to choose between different descriptors in order to keep the listing title below the preset character limit.

The process of listing movie discs has been streamlined by eBay so that the seller need only enter the product UPC code into a prompt, and eBay automatically populates the listing title to include the movie's official title along with the format of the disc and the year that the particular edition was released.[14] This automatically generated listing title serves as something of a default, which the seller can then edit by replacing the existing text, though this is uncommon except when space is a constraint, or adding details. In deciding which details to include in their listing title, sellers are advised by eBay's help page to "Say exactly what the item is even if the title repeats the category name and include details such as brand, product name, size, or artist."[15] Commonly used descriptors include the names of the movie's recognizable star(s), the specific edition (e.g., if it is a special edition), the item's condition (e.g., whether it is still new in the original packaging), and the type of shipping offered.

eBay's search results are quite sensitive to which search terms are used. This sensitivity is due to the operation of eBay's default search, which is "all words any order." This generally requires every word in the search string to appear in the listing title for the listing to appear in search results. For example, on September 27, 2010, a search for *Batman Begins* DVDs using the

---

[14] eBay also automatically populates the product description page of the listing with important product specifications.
[15] http://pages.ebay.com/help/sell/seller-tips.html

string “Batman Begins DVD” returned 699 listings, “Batman Begins 2005 DVD” returned 265 listings, and “Batman Begins on DVD” returned 5 listings. This disparity is due to many titles omitting the year and most titles omitting the word “on.”

To the extent that the wording of the listing title is an important determinant of whether a listing appears in a buyer’s search results, we should reasonably expect professional eBay sellers to include terms that potential buyers are likely to enter into their search strings. Table 2 demonstrates that sellers with more experience, where we use feedback score as a proxy for experience, are more likely to include the word “new,” “disc,” and the disc’s format in an auction’s listing title. While 95 percent (440 of 463) of new items listed by sellers with the highest feedback scores include the word “new,” 15 percent (2 of 13) listed by sellers with the lowest feedback scores do so. Similarly, the inclusion of the word “disc” and the movie’s format is more prevalent among sellers with higher feedback scores. Thus, the inclusion of these descriptors appears to be a skill acquired by the most experienced sellers.[16]

Individual sellers tend to be predictable in their wording choice. The top panel of Figure 2 demonstrates that of the 339 sellers whose DVD listing history includes at least 2 listings for new items, 183 include the word “new” in nearly all (greater than 95 percent) of their listings, 96 include “new” in almost none (5 percent or less) of their listings, leaving only 60 whose listings vary substantially as to the inclusion of “new.” This pattern persists when we look only at sellers with 5 or more listings (the 93rd percentile of sellers by number of listings) and sellers with 10 or more listing (the 97th percentile of sellers), as illustrated in the two bottom-most panels in Figure 2, and also extends to the word “disc” and to the movie format. It follows that whether a listing

[16] We note that the item’s condition and the disc’s format are automatic fields that the seller is prompted by eBay to populate during the listing process (which is then included on the product description page) and thus our detection of these attributes does not depend on the information conveyed in the listing title.

includes certain descriptors can be predicted quite well by the frequency with which the seller included those descriptors in listings for similar items.

The between-seller differences indicated in Figure 2 are not specific to the particular words we have focused on, but reflect differences in sellers' tendencies to include other descriptors as well. This point is illustrated in Figure 3, which shows the percentile rank of sellers as to the median number of words included in the seller's listings. In the top panel are the 96 sellers that include the word "new" in fewer than 5 percent or less of their listings. The 50th percentile of these sellers includes 6.0 words in the median listing. In the bottom panel are the 183 sellers that include "new" in greater than 95 percent of their listings. The 50th percentile of these sellers includes 8.0 words in the median listing. Clearly, the tendency is for the latter group to include more total words in their listing titles as the difference is large enough to not be explained by the difference as to the inclusion of "new" alone. These patterns underscore a fundamental difference in sellers' approach to listing, which we interpret as reflecting varying levels of sophistication.

Among sellers whose inclusion of "new," "disc" or format varies substantially across listings, we find that the title of the movie explains a significant fraction of this variation. In particular, movies with longer titles (e.g., "Harold & Kumar Escape from Guantanamo Bay," Die Hard 4: Live Free or Die Hard," "Miss Pettigrew Lives for a Day," "Pirates of the Caribbean: At World's End," and "The Scorpion King 2: Rise of a Warrior") are less likely to have these descriptors included in the listing title, which, at the time of our data collection, was required to be no longer than 55 characters.[17] Figure 4 shows that among the 60 sellers for whom the inclusion of "new" varies by listing, the likelihood that the seller includes both the word "new" and the disc's format is decreasing in the length of the movie title. Thus, many sellers who would

[17] In July 2011, eBay increased the character limit to 80.

otherwise include both descriptors in the listing title deviate from this practice when the character limit is binding. The implication is that when the seller's own listing history is insufficient for predicting the inclusion of key descriptors, the movie title fills in part of the gap.

4. THEORETICAL MODEL

We model a market in which sellers compete by listing auctions on a common platform.[18] The auctions are conducted simultaneously and potential buyers ("buyers") can participate in multiple auctions simultaneously as in Peters and Severinov (2006). We depart from Peters and Severinov (2006) by assuming that buyers are not ex-ante aware of all competing auctions and incur search costs in identifying them, analogous to the price dispersion literature.[19] In this way, our model can be thought of as combining an emerging literature on simultaneous auctions with the established literature on price dispersion in fixed-price settings. In what follows, we describe the model, establish the equilibrium strategies, and provide comparative static results.

A) Model primitives

Consider a market for a homogenous good comprised of $m \geq 2$ buyers, each with unit demand. Buyers' valuations, v, are private information, independent and identically distributed according to known distribution F, which has full support over the grid $\Omega = \{v_0, v_0 + d, \ldots, v_0 + Td\}$, for some step size $d > 0$ and positive integer T. Buyers are risk neutral so that a buyer with valuation v who obtains a single unit at price p receives surplus $v - p$.

---

[18] While an important aspect of the eBay environment, our analysis does not account for posted-price BIN listings. Our objective is to explain price dispersion in auctions, whereas price dispersion in posted prices has been modeled extensively. Further, within our formulation, the presence of a posted-price alternative serves only as a bound on auction prices and hence price dispersion among auctions, but adds little to the understanding of why prices in auction markets may be dispersed in the first place.

[19] The analogy requires that what is uncertain to buyers in both auction and fixed-price settings is the expected price the buyer will pay, conditional on searching.

The supply side consists of $n = 2$ sellers offering the good for sale via dynamic second-price auction, which is described in detail below.[20] Upon arriving in the market and deciding to list an item for sale, seller $j \in \{1,2\}$ chooses a starting price, S_j. Sellers make their listing decisions simultaneously and do not observe the listing decisions of the competing seller. We assume sellers vary as to their alternative use value, w_j, which are private information, independent and identically distributed according to G, which has full support over the grid Ω.

Frictions arise as auctions differ as to their visibility to buyers. Consider some listing technology, z_j, such that $z_j = 1$ makes auction j visible to all buyers ("high visibility"), whereas $z_j = 0$ makes the auction visible only to buyers who incur a costly search ("low visibility"). The listing technology is available to all sellers, though not all sellers are informed of how to use it. We therefore distinguish between *informed* sellers, for whom $z_j = 1$, and *uninformed* sellers, for whom $z_j = 0$. Let $\theta \in (0,1)$ denote the probability that a given seller is informed. Informed sellers know that they are informed, but do not know whether or not the opposing seller is. Uninformed sellers, in contrast, are unaware of the distinction between sellers and assume that $z = 1$ for all sellers including themselves.[21]

Buyers can mitigate the above friction by conducting a broad search. In doing so the buyer incurs cost, c_i, which is idiosyncratic, assumed independent across i and identically distributed according to H, which has full support and no mass points on the interval $[0, C]$. We assume $C \geq d$, so that the cost of search is nontrivial. Our primary focus is on cases where the

[20] The model in Peters and Severinov (2006) permits an arbitrary number of simultaneous auctions. We restrict attention to two auctions for simplification as the number of cases to consider increases exponentially with the number of sellers. Our results extend to a richer model with more than two sellers upon making the appropriate modifications.

[21] It seems reasonable to assume that uninformed sellers believe their auctions to be high visibility. If not, then they would conceivably seek out the technology necessary to make their auction high visibility. An alternative modeling approach might allow sellers to obtain the technology for a cost (i.e., the cost of learning), with the only substantive difference between this model and ours being whether or not sellers of low-visibility auctions know that their auctions are low visibility. For our results to hold, it is only necessary that informed sellers know they are informed.

auctions differ as to their visibility. In these cases, we refer to buyers who search broadly as "aware" and those who search narrowly as "unaware."

The search process we consider most closely parallels models of fixed-sample search within the price dispersion literature (e.g., Burdett and Judd 1983). In these models, buyers select the number of price quotes they wish to receive, while the set of sellers sampled is effectively random. Within the current setting, a narrow search determines the number of auctions a buyer will be aware of, *conditional* on the visibility of the auctions: zero if $(z_1, z_2) = (0,0)$; one if $max\{z_1, z_2\} = 1 > 0 = min\{z_1, z_2\}$; and two if $(z_1, z_2) = (1,1)$, while searching broadly necessarily returns both auctions. A second type of search considered in the literature is a sequential search in which a buyer bases the decision to sample an additional price quote on those already received. De los Santos, Hortacsu and Wildenbeest (2012) examine search behavior for purchasers of books online and reject the sequential search paradigm in favor of fixed-sample search. Given the obvious parallels to the current setting, we feel that the fixed-sample search is appropriate.

The game proceeds as follows. Buyers arrive simultaneously and make their search decisions, after which they learn their valuations and receive an index $i = 1, \ldots, m$. The bidding then proceeds in a sequential manner. As in Peters and Severinov (2006), buyers are given an opportunity, in order of their index, to bid in one of the auctions or pass. Bids must be at least as high as the starting price if no bids have been received and the standing price plus an increment, $e < d$, otherwise.[22] After each bid, the standing price and identity of the high bidder are updated, and the buyer has the opportunity to bid again-- not necessarily in the same auction. After buyer

[22] The minimum bid increment, *e*, is set by eBay (equal to 25 cents for standing prices between \$1.00 and \$4.99 and 50 cents for standing prices between \$5.00 and \$24.99), to be distinguished from a step *d*, a modeling assumption used to construct the grid of valuations. Following Peters and Severinov (2006), we consider an equilibrium in which bids are increased by at least $d > e$, so that *e* does not constrain a buyer's choice of bid.

i passes, each previous bidder is given the opportunity, in order of index, to submit a new bid (in either auction) or to pass. Once all of these buyers have passed, buyer $i + 1$ is given the opportunity to bid. Bidding continues until all buyers pass at which point, the high bidder in each auction obtains the item at the final standing price, which constitutes the ending price.[23]

B) Equilibrium price dispersion

We seek to establish a perfect Bayesian equilibrium of the game consisting of sellers' starting prices, buyers' search decisions and bidding strategies as well as a set of beliefs that rationalize these strategies. In solving for the equilibrium, we work backwards beginning with the bidding process.

As of the start of the bidding, buyers' information sets consist of the starting prices in A_i, the set of auctions which they have identified from their search, as well as beliefs over the number of buyers who are aware of each auction in A_i. Following Peters and Severinov (2006), we establish a bidding strategy in which prices increase incrementally. The bidding strategy is formalized as follows.

Definition 1: *When buyer i's index is called, the bidding strategy, $\beta^* \equiv \beta^*(v_i, A_i)$, calls for bidder i to do the following:*

(a) *If buyer i is the current high bidder in any auction, or if v_i is less than or equal to the lowest standing price in A_i, the buyer passes.*

[23] Additional details that correspond to eBay exactly are: In determining the standing price, the second-highest bid in an auction refers to the second-highest bid received by a distinct bidder and if two or more bidders submit the same high bid, the first submitter is the high bidder. A condition in Peters and Severinov (2006) and our model that differs from eBay but greatly simplifies the analysis is that the standing price is equal to the current second-highest bid when at least two bidders are present. On eBay, the standing price is the second-highest bid plus the minimum bid increment, e. This simplification is equivalent to assuming that bidders only consider bid increases of at least one dollar, perhaps because bidders incur a small time/effort cost to placing a bid that makes it not worthwhile to bid in very small increments.

(b) *Otherwise, if there is a unique lowest standing price in* A_i*, buyer* i *bids in this auction. The bid amount is the smallest value on the grid above the standing price.*

(c) *If multiple auctions in* A_i*, have the same standing price, buyer* i *bids in an auction that has either not received bids or in which the standing price has increased since the last change in the identity of the high bidder. If multiple auctions satisfy these criteria or if neither auction does, the buyer bids in each auction* $j \in \{a \in A_i | z_a \in min\{z_1, z_2\}\}$ *with equal probability. The bid amount is the smallest value on the grid above the standing price.*

Parts (a) and (b) are straightforward extensions of the strategy in Peters and Severinov (2006), where we have allowed for unaware bidders. The rationale for part (c) is as follows. In choosing between two auctions with the same standing price, a buyer prefers the auction where she is more likely to become the high bidder. The inference is made over the current high bid, which is unobserved. The logic is explained in comprehensive detail in Peters and Severinov (2006) (pages 227-228). The upshot is that a buyer can infer that the high bid in an auction in which the standing price has changed since the last change in high bidder, is equal to the current standing price. In contrast, the high bid in an auction in which the standing price has not changed since the last change in high bidder is d greater than the standing price. Therefore, only in the former auction will his bid make him the high bidder. Unique to the current setup is the requirement that when indifferent, a buyer chooses an auction j such that $z_j \in min\{z_1, z_2\}$. This condition acknowledges that unaware buyers bid only in auctions with high visibility. An aware buyer therefore prefers to be the high bidder in the auction where there is less chance she will be outbid, particularly by an unaware buyer.

Following bidding strategy β^*, standing prices increase incrementally by d until a point is reached at which all non-winning bidders drop out of the bidding. This occurs when the lowest standing price among the auctions they are aware of is equal to their valuation. Peters and Severinov (2006) refer to this strategy as "efficient bidding," as it allocates the items to the bidders with the highest valuations, so that even ex-post, there is no incentive to any buyer to change his bidding strategy.

Lemma 1: *The bidding strategy β^* is a perfect Bayesian equilibrium of the bidding subgame.*

We now consider the search strategy. Upon entry, each buyer decides whether or not to search broadly, given equilibrium beliefs over the strategies of other buyers. Under a general set of conditions, buyers will separate themselves such that buyers with low costs will search broadly and those with high search costs will search narrowly. Formally, let $\Pi(1;c)$ denote a buyer's expected surplus from conducting a broad search, given beliefs that buyers for whom $c_i \leq c$ search broadly, and given beliefs over starting prices. Let $\Pi(0;c)$ be defined analogously for a buyer contemplating a narrow search. Define c^* as the indifferent bidder type satisfying:

[1] $$c^* = \Pi(1;c^*) - \Pi(0;c^*),$$

where a solution exists, and $c^* = C$ when $C < \Pi(1;C) - \Pi(0;C)$. When equation [1] is satisfied for some $c^* \in (0,C)$, then buyers will separate themselves by the type of search undertaken. If not, then $c^* = C$ and all buyers search broadly.

The following proposition characterizes the conditions giving rise to price dispersion. The results generalize to an arbitrary distribution of starting prices, so it is not specified in the proposition. First, some additional notation is needed. Let P_1 and P_2 denote the ending prices in auctions 1 and 2, respectively, conditional on a sale in that auction. Suppose the two auctions differ as to their visibility and assume in what follows that auction 1 is the high visibility auction and auction 2 the low. Let V_u^2 and V_a^2 denote the 2^{nd} highest valuation of unaware and aware buyers, respectively.[24] The value $M_1 = max\{V_u^2, S_1\}$ constitutes a lower bound on P_1, since this would be the ending price in auction 1 were there to be no aware buyers bidding in it. Similarly, let $M_2 = max\{V_a^2, S_2\}$, which constitutes a lower bound on P_2 only if auction 1 is won by an unaware buyer. We now express our central result.

Proposition 1: *It is a perfect-Bayesian equilibrium for buyers to search broadly if $c_i \leq c^*$ and bid according to β^*. In such an equilibrium,*

(a) *If $z_1 > z_2$ and auction 2 results in a sale, P_2 will be no greater than: i) P_1 if auction 1 results in a sale; or ii) S_1 if it does not. Conditional on both auctions resulting in a sale and $z_1 > z_2$, P_1 strictly exceeds P_2 if $M_1 > M_2$.*

(b) $E[|P_1 - P_2|] > 0$ *if and only if $z_1 \neq z_2$.*

(c) $E[|P_1 - P_2|]$ *is strictly decreasing in the proportion of buyers that are aware.*

In the equilibrium described in Proposition 1, the ending prices in the two auctions are equal as long as no unaware buyer is among the two highest bidders in either auction. Part (a) of the proposition specifies the circumstances when this is not the case. Price dispersion (within this equilibrium) requires auction 1 to be won by an unaware buyer. Two assumptions, $\theta < 1$ and $C \geq d$, are sufficient to guarantee that a given buyer is unaware with positive probability so long as $z_1 \neq z_2$.[25] That is part (b) of the proposition. By demonstrating a monotonic relationship between the proportion of aware buyers and price dispersion, part (c) differentiates our model from models of price dispersion in fixed-price sales as illustrated in Figure 1. Stahl (1989) showed that Bertrand competition and monopoly pricing are limiting cases when all buyers are

[24] In the case of ties, the rank order increases by one for every buyer. So if there are two buyers with the same highest valuation, then that value constitutes the first and second highest valuations.

[25] The former insures that the expected benefit of a broad search is positive and the latter insures that the cost is non-negligible.

aware and unaware respectively, so that price dispersion is at its greatest somewhere between the two extremes.[26] Brown and Goolsbee (2002) find empirical support for this prediction.

We noted that the equilibrium starting price strategies are not essential to explaining price dispersion as in Proposition 1. However, the starting price offers the informed seller an additional strategic variable with which to influence the price. In fact, the following proposition demonstrates that visibility differences create an incentive for informed sellers to increase their starting price, relative to what it would have been were the seller uninformed, thereby exacerbating the effect of unaware buyers on prices.

Proposition 2: *There exists a symmetric starting price strategy, $\sigma^* \equiv \sigma^*(w_j, z_j)$, that maximizes the surplus of seller j, with valuation w_j and visibility z_j, given opposing sellers set starting prices according to σ^* and given equilibrium behavior by buyers, such that $\sigma^*(w, 1) \geq \sigma^*(w, 0) \geq w$ and $\sigma^*(w, 1) > \sigma^*(w, 0)$ under certain parameterizations for certain values of w.*

The proposition shows that an informed seller sets a (weakly) higher starting price than if she were uninformed. Peters and Severinov (2006) showed that competition from other sellers puts downward pressure on the equilibrium starting price, relative to its level under monopoly (e.g., Myerson 1981, Riley and Samuelson 1981). However, greater visibility confers market power to the seller who is effectively a monopolist over the unaware buyers, allowing the seller to set a higher starting price.

[26] Baye, Morgan and Scholten (2006) generalize this result to a broader range of models.

5. EFFECT OF LISTING WORD DIFFERENCES ON PRICE DISPERSION

We now test the main theoretical prediction, which is that price dispersion decreases in the fraction of buyers who are aware. Since we do not know directly which auctions a buyer is aware of, we use differences in the inclusion of key words as a proxy for the fraction of unaware buyers. We also provide supporting evidence from tests of the other predictions of the model.

A) Theoretical predictions

The theoretical model explains price dispersion as arising from frictions which lead buyers to become unaware of certain auctions and consequently bid up the price they are aware of beyond what they could have paid in another auction they are not aware of. In particular, buyers that enter additional modifiers in the search bar will not observe auctions that do not include each modifier in its listing title.[27] Other frictions besides listing-wording differences may also be important. For example, if the auctions under consideration have different start times or end times, there may be buyers in either auction that are not aware of the competing auction due to when they entered the market. Additionally, if bidding itself is costly, buyers may place only a limited number of bids. Anticipating that a given bid may be their last, a buyer would reasonably bid above the minimum increment above the standing price, causing the buyer to overpay relative to the competing auction. While likely important, these mechanisms operate independent of the mechanism of interest and only serve to add noise to the estimate of the effect of listing-wording differences on price dispersion.

[27] Wording differences may affect search in other ways. First, given the large number of listings that appear in search results, bidders may favor listings with titles that directly indicate the desired item even if listings with and without the words both appear in the search results. Second, some words may reflect higher quality in a way that is observed (or inferred) by the bidder but not observed by the researcher (e.g., sellers with more descriptive titles may be more reliable). These explanations alter the mechanism through which differences in listing wording translate to differences in price, but ultimately do not change our conclusion that more sophisticated sellers choose the wording of their listings to maximize revenue.

To formalize our primary prediction, let $z_j^w = 1$ if auction j includes word w that some buyer include in their search string and $z_j^w = 0$ otherwise. The following set of predictions apply to pairs of auctions consisting of auctions $j = 1,2$. All notation is as in Section 4.

Prediction 1: *There exists some key word, w, such that $E[|P_1 - P_2|]$ is increasing in $|z_1^w - z_2^w|$, holding constant $\left|z_1^k - z_2^k\right|$ for all other words, k.*

Prediction 1 uses wording differences to proxy for the fraction of unaware buyers. We cannot directly verify the relationship between listing wording and the fraction of unaware buyers since we do not observe buyers' search results. Nevertheless, if wording differences do increase the fraction of unaware buyers, then the following is necessarily true.

Prediction 2: *If $z_1^w > z_2^w$, then $E[P_1] - E[P_2] > 0$.*

The mechanism through which wording differences lead to greater price dispersion works through two channels: (1) By causing bidders who enter key descriptors into their search strings to become unaware of auctions that do not include these terms in their listing title. This leads them to sometimes bid in the auction with the higher standing price so that prices become more dispersed. (2) Informed sellers take this behavior into account and set a higher starting price along with including key descriptors in the listing title. Since our measure of price dispersion conditions on the auctions transacting, a standard result from the optimal auctions literature is that this leads to a higher final price. These two channels generate the following two predictions.

Prediction 3: *If $z_1^w > z_2^w$, then unaware buyers may bid in auction 1 when auction 2 has the lower standing price of the two.*

Prediction 4: *If $z_1^w > z_2^w$, then $E[S_1] > E[S_2]$.*

B) Test of Prediction 1

i) Specification

Our analysis measures price dispersion over pairs of auctions for identical items sold by similarly reputable sellers and whose end times are in close proximity. We begin by grouping all auctions for the same movie title that are of the same format (DVD or Blu-ray), edition, and condition (new or used) and that are offered via the same shipping method (priority or standard).[28] Next, we construct pairs from all auctions conducted by distinct sellers and whose ending times were no greater than twelve hours apart. This constitutes a rolling twelve-hour window, which relies on each auction's ending time as the upper bound for one window and a lower bound for another. The twelve-hour cutoff is appropriate as it is a sufficiently narrow window that a buyer should reasonably be indifferent between purchasing at any one auction within the set.[29] When more than two auctions within the same group have ended within the same twelve-hour period, each auction appears in multiple pairs. For example, in the case of three auctions *A*, *B* and *C*, we construct the following three pairs: (*A*, *B*), (*A*, *C*) and (*B*, *C*). To account for the nonstandard error structure that arises in the pairwise analysis, we compute standard errors using a nonparametric bootstrap method.

Since sellers differ as to their reliability, we restrict our sample to pairs of auctions sold by sellers with similar feedback scores. Livingston (2005) demonstrates that returns to additional positive feedback reports are steeply declining beyond the first 25 reports. With this result as our guide, we constructed the following bins within which the feedback ratings of the two sellers in

[28] Several titles include both a widescreen and a full-screen edition. As eBay did not (at the time of data collection) include descriptors for widescreen and full-screen editions as automatic fields, we could only identify these editions by the listing title, which in some cases did not indicate the edition. In these cases, we included all of these editions, including those that were identified as widescreen or full-screen into one group. We have reproduced our results under several sets of assumptions regarding these unknown editions, including dropping all observations of title-format combinations where there is uncertainty over widescreen versus full screen edition.

[29] All of the results of this section have been replicated with a narrower three hour window. As a narrower window leads to a smaller dataset, certain results have a lower level of significance with a three hour window.

the pair (should they be distinct sellers) must fall within: 1-5, 6-35, 36-215, 216-1,291, 1,292-7,766, 7,767-46,655 and 46,656+.[30,31]

Since pairs contain information that would not be evident from larger groups, the unit of analysis is the pair. Were we to construct groups from three or more auctions (using a measure of price dispersion that incorporates all prices, such as the COV), our measure of price dispersion would not fully reflect wording differences in the group if these differences were exhibited in only a subset of auctions. The pairwise approach therefore offers a more efficient estimation (i.e., a group-level measure of wording difference is not a sufficient statistic for pair-level measure).

Our specification captures the essential elements of the approach used in much of the price dispersion literature (Baron, Taylor and Umbeck, 2004; Brown and Goolsbee, 2002; and Lewis, 2008). These papers calculate price dispersion using deviations from the price predicted from product and seller characteristics as well as time fixed effects. Similarly, we calculate price dispersion using pairwise differences in price, where both auctions in the pair have the same predicted price based on product and seller characteristics.

We estimate the following model, where the subscript p indicates the pair,

[2] $$E\left[D_p|W_p\right] = \delta_0 + Z_p\delta_1 + \varepsilon_p.$$

Equation [2] expresses the expected absolute deviation in prices, $D_p \equiv ln\left(\left|P_i - P_j\right|\right)$, as a function of $Z_p = \left(\left|z_1^{new} - z_2^{new}\right|, \left|z_1^{disc} - z_2^{disc}\right|, \left|z_1^{format} - z_2^{format}\right|\right)$.[32] While the terms "new," "disc," and the disc's format are not nearly exhaustive of the descriptors used in listings,

[30] The cutoffs correspond to a log base six scale, but the results are robust to alternatives. We chose this functional form to roughly match the steep experience slope in Livingston (2005).
[31] We drop the small number of sellers who received less than 99 percent positive feedback, since they generally do not contain matches in the data set for constructing auction pairs.
[32] The log specification provides a better fit of the model. This makes sense theoretically if the variance in valuations (or in prices) is increasing in the mean valuation (price).

they are among the most widely used and have the benefit of not being specific to a particular movie title. Table 3 summarizes the data used in the analysis.

A concern with estimating equation [2] is that within-seller differences in wording may reflect differences in products that are not observable to the researcher (us) but are to bidders. For example, if the inclusion of "new" were correlated with higher quality, then the coefficient on the pair-wise difference in "new" would be biased upward. We address this concern by instrumenting for wording differences.

To identify instruments, we consider the Section-3 results that showed that variation in the inclusion of key words in the listing title is often idiosyncratic to the seller: Many sellers either always or never include the word "new" for new items. Since these sellers do not vary wording across listings, these sellers do not appear to be strategically altering their use of these key words across similar but distinct products. For sellers who vary whether they include certain key words, the length of the movie title is a significant source of variation: *Listing* titles for movies with longer *movie* titles are less likely to include key words due to eBay's listing-title length limit of 55 characters. Since the movie-title length is unlikely to be related to price dispersion of that movie except through its effect on listing-title wording, we use movie-title length as an instrument for listing-wording differences in the auction pair.

We estimate equation [2] via two-stage least squares (2SLS). In the first stage, we estimate the probability that the two auctions in the pair differ as to the inclusion of "new," "disc," and movie format in the listing title. Formally, we have the following stage-1 equations,

$$E\left[Z_p^{new}\right] = \gamma_0^n + F_p\gamma_1^n + M_p\gamma_2^n + I_p\gamma_3^n + \varepsilon_p$$

$$E\left[Z_p^{disc}\right] = \gamma_0^d + F_p\gamma_1^d + M_p\gamma_2^d + I_p\gamma_3^d + \varepsilon_p, \qquad [3]$$

$$E\left[Z_p^{format}\right] = \gamma_0^f + F_p\gamma_1^f + M_p\gamma_2^f + I_p\gamma_3^f + \varepsilon_p,$$

where Z_p as in [2], F_p is a vector of cross-seller listing differences, M_p is the exponential of the number of characters in the movie title (which is the same for both auctions in the pair) and I_p is a vector of item characteristics (which is the same for both auctions in the pair), namely whether the item is new and whether it is a Blu-ray disc.[33] The elements of $F_p = \left(F_p^{new}, F_p^{disc}, F_p^{format}\right)$ correspond to cross-seller differences as to the frequency in which they include "new," "disc," and format in their listing history.

In constructing F_p^{new}, we first calculate the proportion of the listings in each seller's history (not including the listing in question) of new items that include the word "new" in the listing title. A seller whose listings typically include the word "new" should be more likely to include "new" in the listing in question. Taking the within-pair difference between the proportions of the two sellers' listings that include the word "new" therefore provides an estimate of the likelihood that the two auctions in the pair will differ as to the word "new." F_p^{disc} and F_p^{format} are calculated in an parallel manner, except that F_p^{disc} is calculated over the sample of the sellers' other Blu-ray listings and F_p^{format} is calculated over the sample of *all* of the sellers' other listings, regardless of condition or format. Sellers with insufficient history with which to calculate F_p are dropped from the sample.[34]

ii) <u>Results</u>

Column (3) of Table 4 presents results of the 2SLS estimation, while OLS estimates are presented in column (1). The within-pair variation as to the inclusion of "new" that is predicted by equation [3] leads to a 40.2 percent increase in price dispersion. This effect is not statistically significant, however, due to the large standard errors (p = 0.16) relative to column (1). The effect

[33] The exponential form reflects that the length of the listing title is constrained primarily for movies with long titles.
[34] To the extent that the listings of these less professional sellers differ from other sellers beyond simply the wording of the listing title, this step reduces noise in our estimates.

of within-pair variation as to the inclusion of disc is positive and significant ($p < 0.10$), leading to a 96.4 percent increase in price dispersion. Variation as to the inclusion of the format has virtually no effect. A chi-squared test on the three variables as a group produces a p value of 0.21, leaving some doubt that the observed impact of these wording differences is due to chance.

The chi-squared statistic can be interpreted as reflecting the importance of wording difference as a friction, relative to other frictions not tested by the model. To the extent that other frictions, two of which were discussed in Section 5A, are significant, the marginal effect of wording differences is estimated with significant noise. To overcome this problem, the three tests that follow focus only on pairs of auctions with listing-wording differences. All three tests show that the effect of listing-wording differences works in the direction predicted by the theory.

C) Prediction 2

To the extent that differences in listing titles as to the inclusion of the words “new” and “disc” contribute to greater price dispersion, Prediction 2 indicates that auctions whose listing titles include these words should have higher ending prices than those that do not. To test for the effect of “new,” we restrict attention to auction pairs for new items and in which one auction includes the word “new” in the listing title while the other does not. The tests for “disc” and the disc’s format are analogous but with the former restricting attention to auctions for Blu-ray discs, while the latter includes the full sample regardless of condition or format.

Results are summarized in Table 5, demonstrating that auctions that include in their listing titles “new,” “disc” and the movie’s format outperform their counterparts by an average of \$0.53 ($p < 0.05$), \$1.00 ($p < 0.01$) and \$0.30 ($p = 0.14$), respectively. Further, these auctions are more likely to have higher final prices. In identifying auctions with higher prices, we limit attention to price differences of at least \$0.50, which represents the smallest increment by which a bid must exceed the previous high bid.[35] Auctions that include the word "new" have a strictly higher price than the paired auction that does not include "new" in 48 percent of pairs versus 39 percent for the reverse. This difference becomes more stark when we consider larger price differences: auctions that include "new" are 13 percent more likely to outperform their counterpart by at least one dollar than auctions that do not include "new" ($p < 0.10$). A similar pattern holds for pairs that differ as to the inclusion of "disc" and the movie's format.

D) Prediction 3

To the extent that listing-wording differences cause buyers to become unaware, Prediction 3 states that they will be unaware of the auction that does not include key descriptors in its listing title when the competing auction does. We test this prediction by examining the fraction of buyers who place their first bid in the auction within the pair that has a higher standing price when the two auctions differ as to the inclusion of a descriptor such as "new."[36] Bidding in the auction with the higher standing price is indicative of an unaware buyer, so buyers should be more likely to do so in auctions that are more visible due to the inclusion of key words.

The results of the test are in Figure 5. The top panel demonstrates that a greater fraction of unaware buyers bid in the auction that includes the word "new" and "disc" than in the paired auction that does not included these terms. When we judge unaware bidding by standing price differences of at least one dollar, the effect of "new" becomes more pronounced while the effect of "disc" is unaffected. Thus, differences in listing titles as to the inclusion of these terms appear to contribute significantly to unaware bidding. The effect of the format (in both panels) is

[35] Given that each bid must be at least $0.50 higher than the standing price, an buyer following the equilibrium bidding strategy of β^* may still pay a price upwards of $0.50 higher than in the paired auction.

[36] We focus on the bidder's first bid as there may be reasons beyond listing-wording differences that explain why a buyer would place a bid in the same auctions as his last bid.

actually the reverse of the other two terms, however, differences in listing titles as to the inclusion of format was shown to have a negligible effect on price dispersion (Table 4).

E) Prediction 4

Prediction 4 states that key descriptors are included in an auctions listing title by informed sellers who also set higher starting prices. We test this prediction by analyzing pairs of auctions where one auction in the pair includes a particular descriptor such as “new” in its listing title, but the other auction does not. The outcome variable is the *effective starting price*, equal to the sum of the actual starting price and the shipping charge. The effective starting price, as opposed to actual starting price, reflects the true minimum amount a buyer pays upon winning the auction and reflects the fact that an increase in the actual starting price of $1 that is offset by a corresponding decrease in the shipping charge of $1 should have no effect on auction outcomes.[37] Figure 6 indicates that the mean effective starting price among auctions that include the words “new” and “disc” in their listing titles is $1.44 and $3.06 higher, respectively, than in the paired auction that does not ($p < 0.05$). The $0.62 difference in effective starting price explained by differences in the inclusion of movie format is not statistically significant.

6. CONCLUSION

Auction markets offer a number of interesting features with which to derive and test theories of price dispersion. We have proposed a model in which auctions differ as to the ease in which potential buyers can identify various auctions within a set of search results. Price dispersion results as buyers who are unaware of all auctions taking place concurrently bid more than what would have been required to obtain the item in a competing auction. The model offers a set of empirical predictions, which we test with data of movie DVD auctions on eBay.

[37] Hossain and Morgan (2006) make this point.

Our primary result tests a theoretical prediction regarding the fraction of aware bidders on price dispersion. We use differences in the wording of auction listings as a proxy for the fraction of unaware bidders. Specifically, buyers that include additional modifiers in their search strings will be unaware of auctions for the desired item whose listing titles do not include each modifier. Consistent with the theory, we find that differences as to the inclusion of the words "new" and "disc" lead to greater price dispersion.

Next, we test three additional predictions of the theory, the results of which lend additional support to our primary test. First, we find that the inclusion of the words "new," "disc" and format raises the sellers revenue by $0.53 and $1.00 and $0.30, respectively, above auctions for the identical item ending on the same day that do not include those words. Next, we find that the inclusion of the words "new," "disc" and format increase the likelihood that a buyer submits his first bid in that auction when the auction in question does not have the lowest standing price. And third, we find that sellers take this behavior into account in their choice of starting price, such that auctions that include the word "new" in their listing titles have higher starting prices than otherwise identical auctions by an average of $1.44.

Collectively, these results demonstrate the importance of listing wording in explaining price dispersion on eBay. However, this need not be the only friction and that we have not attempted to explain all relevant features of bidding on eBay. Beginning with the paper by Alvin Roth and Axel Ockenfels in 2002, explaining bidding behavior in online auctions has been a subject of much study. Most of these studies however, have treated auctions in isolation. A number of recent papers incorporate competing auctions and dynamic considerations into models of bidding behavior online, but there is still much to be learned.

REFERENCES

Anwar, S., McMillan, R., and M. Zheng (2006) "Bidding Behavior in Competing Auctions: Evidence from eBay," European Economic Review, 50(2), 307-322.

Barron, J.M., Taylor, B.A., and J.R. Umbeck (2004) "Number of Sellers, Average Prices, and Price Dispersion," International Journal of Industrial Organization 22(8), 1041-1066.

Baye, M.R., and J. Morgan (2001) "Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets," American Economic Review, 91(3), 454-474.

Baye, M.R., Morgan, J. and P. Scholten (2004) "Price Dispersion in the Small and Large: Evidence from an Internet Price Comparison Site," Journal of Industrial Economics, 52(4), 463-496.

Baye, M.R., Morgan, J. and P. Scholten (2006) "Information, Search and Price Dispersion," *Forthcoming in Handbook on Economics and Information Systems*.

Brown, J., and A. Goolsbee (2002) "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry," Journal of Political Economy, 110 (3), 481-507.

Burdett, K. and K. Judd (1983) "Equilibrium Price Dispersion," Econometrica, 51(4), 955-969.

Carlson, J. A. and R.P. McAfee (1983) "Discrete Equilibrium Price Dispersion," Journal of Political Economy, 91(3), 480-493.

Chiao, L. and J. Pate (2010) "Internet Auctions and Frictionless Commerce: Evidence from the Retail Gift Card Market," Review of Industrial Organization, 36, 295-304.

Clay, K., Krishnan, R., and E. Wolff (2001) "Prices and Price Dispersion on the Web: Evidence from the Online Book Industry," Journal of Industrial Economics, 49(4), 521–540.

De los Santos, B., Hortacsu, A., and M. Wildenbeest (2012) "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior," forthcoming, American Economic Review.

Einav, L., Kuchler, T., Levin, J. and N. Sundaresen (2011) "Learning from Seller Experiments in Online Markets," NBER Working paper No. 17385.

Haruvy, E. and P.T. Popkowski Leszcztc (2010) "Search and Choice in Online Consumer Auctions," Marketing Science, 29(6), 1152-1164.

Hasker, K. and R. Sickles (2010) "eBay in the Economic Literature: Analysis of an Auction Marketplace," Review of Industrial Organization, 37, 3-42.

Hendricks, K., Onur, I. and T. Wiseman (2012) "Last Minute Bidding in Sequential Auctions with Unobserved, Stochastic Entry," Review of Industrial Organization, 40(1), 1-19.

Hossain, T. and J. Morgan (2006) "...Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on eBay," Advances in Economic Analysis & Policy, 6(2), Article 3.

Huang, C.I., Chen, K.P., Chen, J.R. and C.F. Chou (2008) "Bidding Strategies in Internet Parallel Auctions," *working paper, Available at SSRN: http://ssrn.com/abstract=1079943*.

Lewis, M. (2008) "Price Dispersion and Competition with Differentiated Sellers," The Journal of Industrial Economics, 56(3), 654-678.

Livingston, J.A. (2005) "How Valuable is a Good Reputation? A Sample Selection Model of Internet Auctions," The Review of Economics and Statistics, 87(3), 453-465.

MacMinn, R.D. (1980) "Search and Market Equilibrium," Journal of Political Economy, 88(2), 308-327.

McAfee, R.P. (1993) "Mechanism Design by Competing Sellers," Econometrica, 61(6), 1281-1312.

Myerson, R.B. (1981) "Optimal Auction Design," Mathematics of Operations Research, 6(1), 55-73.

Narasimhan, C. (1988) "Competitive Promotional Strategies," Journal of Business, 61(4), 427-449.

Perloff, J.M. and S.C. Salop (1985) "Equilibrium with Product Differentiation," The Review of Economic Studies 52, 107-120.

Peters, M. and S. Severinov (2006) "Internet Auctions with Many Traders," Journal of Economic Theory, 130, 220-245.

Podwol, J. and H. Schneider (2012) "The Bidder's Curse: A Comment on Malmendier and Lee (2011)," *working paper*

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

Reinganum, J. (1979) "A Simple Model of Equilibrium Price Dispersion," Journal of Political Economy, 87(4), 851-58.

Rosenthal, R.W. (1980) "A Model in Which an Increase in the Number of Sellers Leads to a Higher Price," Econometrica, 48(6), 1575-1580.

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

Salop, S. and J. Stiglitz (1977) "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion," Review of Economic Studies, 44(3), 493-510.

Spulber, D. (1995) "Bertrand Competition when Rivals' Costs are Unknown," Journal of Industrial Economics, 43(1), 1-11.

Stahl, D.O. (1989) "Oligopolistic Pricing with Sequential Consumer Search," American Economic Review, 79(4), 700-712.

Varian, H. (1980) "A Model of Sales," The American Economic Review, 70(4), 651-659.

Wolinsky, A. (1988) "Dynamic Markets with Competitive Bidding," Review of Economic Studies, 55(1), 71–84.

Table 1: Price dispersion summary, by title-format-edition

Movie Title (Format)	Edition	N Groups	Mean N Transactions per Group	Mean of Group-wise Mean Price	Mean of Group-wise Mean COV
Batman Begins (DVD)	Regular Edition (1 disc)	6	2.50	7.36	0.18
Batman Begins (Blu-ray)	Regular Edition (1 disc)	20	2.95	21.31	0.08
Camp Rock (DVD)	Regular Edition (1 disc)	28	3.18	13.78	0.10
Camp Rock (Blu-ray)	Regular Edition (1 disc)	4	2.00	12.22	0.06
Casino Royale (DVD)	Regular Edition (2 disc)	6	2.00	10.19	0.16
Casino Royale (Blu-ray)	Regular Edition (1 disc)	8	2.38	19.23	0.07
College Road Trip (DVD)	Regular Edition (1 disc)	7	2.71	9.07	0.14
College Road Trip (Blu-ray)	Regular Edition (1 disc)	3	2.67	13.37	0.13
Harold & Kumar Escape from Guantanamo Bay (Blu-ray)	Special Edition (2 disc)	3	2.33	18.00	0.05
Knocked Up (DVD)	Regular Edition (1 disc)	6	3.67	7.09	0.21
Die Hard 4: Live Free or Die Hard (Blu-ray)	Regular Edition (1 disc)	3	2.00	18.50	0.14
Miss Pettigrew Lives for a Day (DVD)	Regular Edition (1 disc)	4	2.00	11.81	0.09
Pirates of the Caribbean: At World's End (DVD)	Regular Edition (1 disc)	16	3.13	8.30	0.21
Pirates of the Caribbean: At World's End (Blu-ray)	Special Edition (2 disc)	5	2.00	18.76	0.11
Street Kings (DVD)	Regular Edition (1 disc)	9	2.33	8.93	0.14
Street Kings (Blu-ray)	Regular Edition (1 disc)	3	2.33	20.71	0.06
The Bank Job (DVD)	Regular Edition (1 disc)	5	2.20	9.27	0.20
The Notebook (DVD)	Regular Edition (1 disc)	10	2.20	12.03	0.12
The Scorpion King 2: Rise of a Warrior (DVD)	Regular Edition (1 disc)	14	2.71	7.34	0.16
Transformers (DVD)	Regular Edition (1 disc)	9	2.78	9.63	0.18
Transformers (Blu-ray)	Special Edition (2 disc)	40	4.55	25.18	0.08
All Titles		215	3.03	15.11	0.12

Notes: The unit of observation is the group, consisting of all such auctions of the same title, format and edition ending on the same date. The sample consists of all completed auctions for new DVDs and Blu-ray discs offered with standard shipping in groups of at least two. We present only editions with at least 3 groups in the sample. The remaining groups are represented in the All Titles row.

Table 2: Frequency of "new," "disc" and the disc's format by seller feedback score

Seller feedback score	N. Unique Sellers	New			Disc			Format		
		N. Include	N. New Items	% Include	N. Include	N. Blu-ray Discs	% Include	N. Include	N. Total Listings	% Include
0 - 5	41	2	13	15.4	4	7	57.1	45	54	83.3
6 - 35	185	18	55	32.7	21	35	60.0	197	223	88.3
36 - 215	683	102	231	44.2	110	172	64.0	794	901	88.1
216 - 1,295	968	382	620	61.6	305	436	70.0	1,616	1,828	88.4
1,296 - 7,775	563	852	1,198	71.1	313	401	78.1	1,975	2,188	90.3
7,776 - 46,655	140	293	365	80.3	120	176	68.2	749	792	94.6
46,656 +	34	440	463	95.0	20	29	69.0	644	652	98.8
All Listings	2,614	2,089	2,945	70.9	893	1,256	71.1	6,020	6,638	90.7

Notes: The unit of observation is the listing (auction and BIN). The data are from the full set of all auction and BIN listings. Summary statistics for the three descriptors are based on the following subsamples. (1) For "new," the subsample is all listings for new movie discs. (2) For "disc," the subsample is all listings for Blu-ray discs. (3) For format, the subsample is all listings for all movie discs regardless of condition or format. The rows indicate a range of seller feedback ratings. The columns indicate the number of listings with the word, number of listings in the subsample and percentage of listings in the subsample with the word, respectively.

Table 3: Summary statistics for pairwise analysis

	N.	Mean	Std. Dev.	Min	Max
Absolute difference in final price	783	2.21	1.96	0.00	14.26
Log of absolute in final price	783	0.25	1.40	-6.91	2.66
Differ as to the inclusion of "new"'	286	0.42	0.49	0	1
Differ as to the inclusion of "disc"	206	0.24	0.43	0	1
Differ as to the inclusion of format	783	0.12	0.32	0	1

Notes: The unit of observation is the auction pair (N =783). Summary statistics for "Differ as to the inclusion of "new"" and "Differ as to the inclusion of "disc" are presented for only new items (N =286) and Blu-ray discs (N =206), respectively.

Table 4: Estimated model of the effect of wording differences on price dispersion

	(1)	(2)	(3)	
Estimation	OLS	OLS	2SLS	
Δ Include "new"	0.318***	0.167	0.338	*F* =70.48
	[0.111]	[0.155]	[0.241]	
Δ Include "disc"	0.024	0.250	0.675*	*F* =59.76
	[0.247]	[0.350]	[0.409]	
Δ Include the format	0.006	-0.013	0.065	*F* =93.61
	[0.156]	[0.235]	[0.314]	
Constant	0.199***	0.140	0.058	
	[0.060]	[0.097]	[0.122]	
Seller feedback controls	Yes	Yes	Yes	
Drop sellers with no history	No	No	No	
Observations	783	346	346	
R-squared	0.007	0.004		
Chi-squared (p value)	8.44 (0.038)	1.48 (0.687)	4.57 (0.206)	

Notes: The unit of observation is the auction pair. The models in columns (1) - (3) estimate the log of the absolute difference in price via OLS, OLS and 2SLS, respectively. The F statistics corresponding to the first-stage regressions are presented alongside the coefficient estimates in column (3). "Seller feedback controls" indicate the sample is restricted to pairs in which the sellers' feedback scores are no further than 1 unit apart on a log base 6 scale and sellers with less than 99 percent positive feedback are dropped from the sample. "Drop sellers with no history" applies to observations from sellers that lack a sufficient sample for calculating the instruments. Coefficients are reported with *, **, and *** indicating significance at the 10, 5 and 1 percent level, respectively. Reported in brackets are standard errors calculated using a nonparametric bootstrap method.

Table 5: Effect of wording on final price

	"New" (N =120)			"Disc" (N =85)			Format (N =92)		
	Mean price	Price is >$0.50 higher	Price is ≥$1.00 higher	Mean price	Price is >$0.50 higher	Price is ≥$1.00 higher	Mean price	Price is >$0.50 higher	Price is ≥$1.00 higher
Term included	$16.79	0.48	0.43	$15.65	0.52	0.46	$12.01	0.48	0.40
Term not included	$16.26	0.39	0.31	$14.65	0.32	0.26	$11.71	0.36	0.28
Difference	$0.53	0.08	0.13	$1.00	0.20	0.20	$0.30	0.12	0.12
	(p =.04)	(p =.16)	(p =.06)	(p <.01)	(p =.02)	(p =.01)	(p =.14)	(p =.10)	(p =.08)

Notes: The unit of observation is the pair. One sample consists of pairs in which one auction includes the word “new” in its listing title and the other does not (N =120). A second sample consists of pairs in which one auction includes the word “disc” in its listing title and the other does not (N =85). A third sample consists of pairs in which one auction includes the disc’s format in its listing title and the other does not (N =92). P values are derived from a one-sided test and calculated using standard errors from a nonparametric bootstrap estimation.

FIGURES

Figure 1: Comparison of price dispersion (PD) predictions for auction versus fixed-price markets

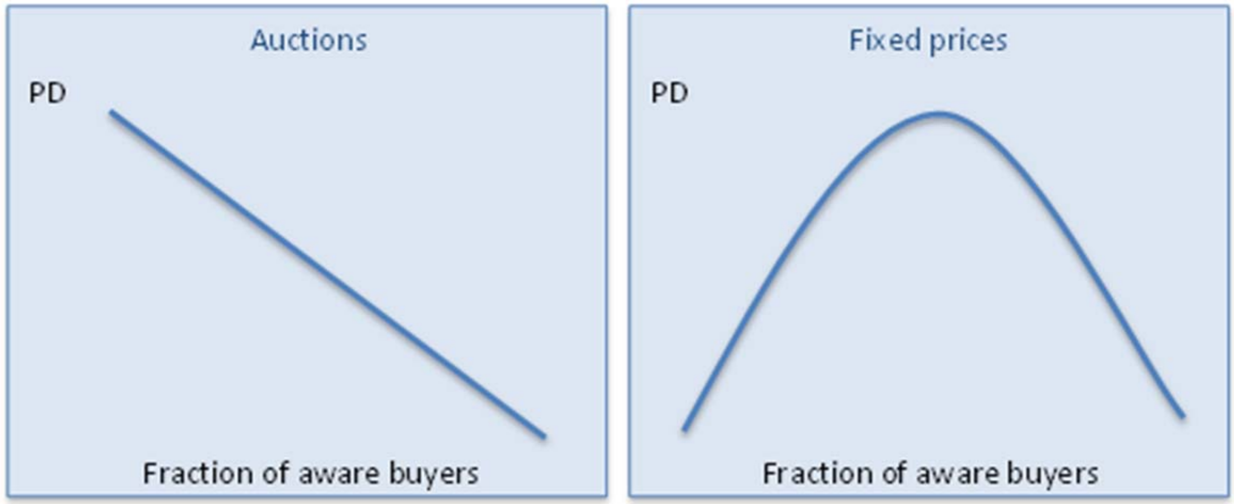
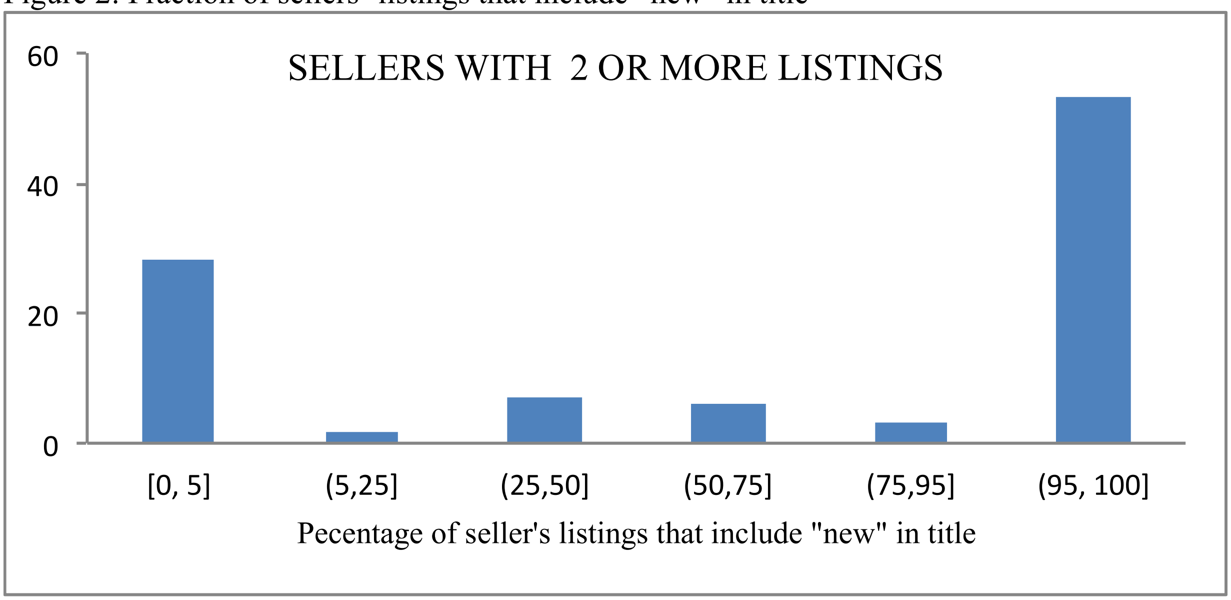
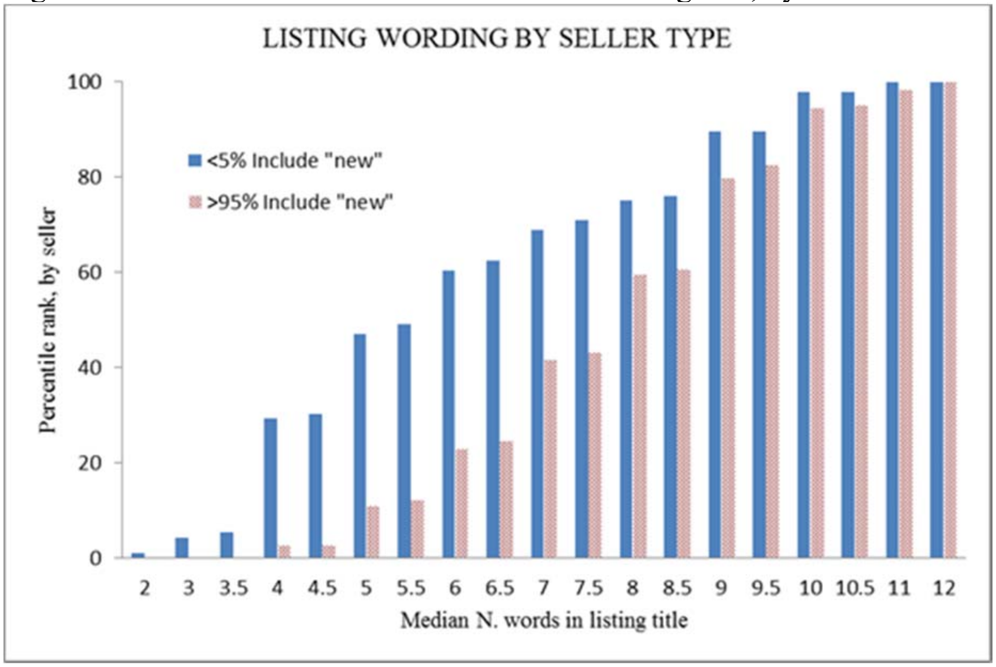


Figure 2: Fraction of sellers' listings that include "new" in title



Notes: The unit of observation is the seller. Bar heights indicate the number of sellers that include the descriptor at the frequency given on the horizontal axis. The sample includes sellers with two or more listings in the data (N =339). Results from samples with sellers with at least 5 (N =113) and 10 (N =46) listings in the data are similar.

Figure 3: Median number of words included in listing title, by seller



Notes: The unit of observation is the seller. Bar heights indicate the percentile rank of sellers based on the median number of words in their listing titles. The two sets of bars are sellers with the word "new" in 5 percent or less of listings, and 96 percent or more of listings, respectively.

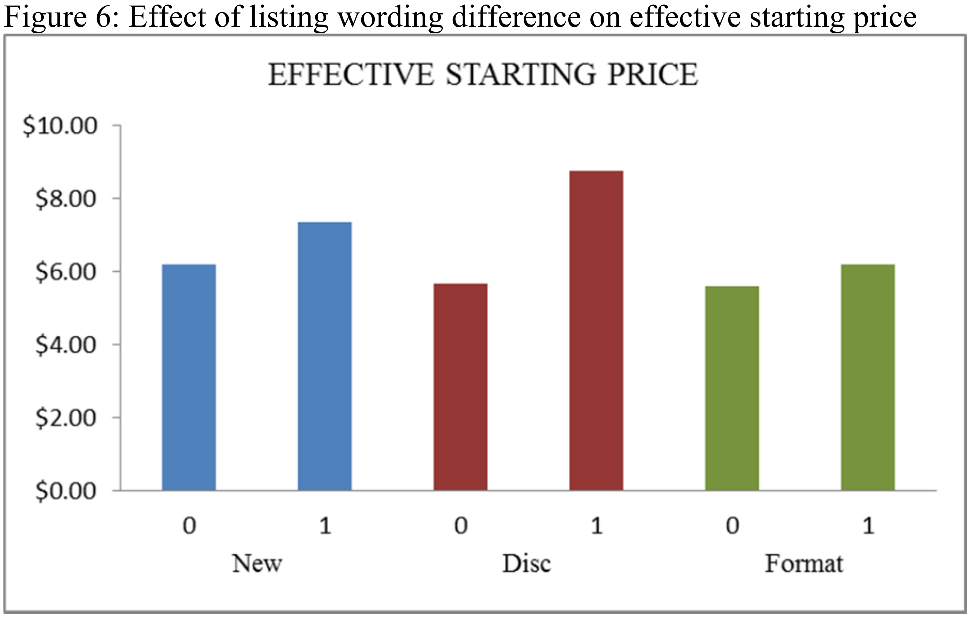
Figure 4: Effect of movie title length on listing wording

Notes: The curves show the estimated probability that the seller includes the word "new" in the listing title, and both "new" and the format in the listing title. Estimates are from OLS where the dependent variables are the number of characters in the movie title and its square and where seller fixed effects are included for the 60 sellers whose listings vary as to the inclusion of the word "new" (N = 558 listings for new items).

Figure 5: Effect of listing wording difference on unaware bidding

Notes: The unit of analysis is the bidder-pair combination. The three subsamples analyzed are: 1) Bidders in pairs of auctions for new items that differ as to the inclusion of "new"; 2) Bidders in pairs of auctions that differ as to the inclusion of "disc"; and 3) Bidders in pairs of auctions that differ as to the inclusion of the disc's format. Bar heights in the left panel indicate fraction of buyers whose first bid in the pair is in the auction with the higher standing price when that auction included the word ($z_j^w = 1$) and among those that do not include the word ($z_j^w = 0$). Bar heights in the right panel indicate fraction of buyers whose first bid in the pair is in the auction with the higher standing price by at least $1 when that auction included the word ($z_j^w = 1$) and among those that do not include the word ($z_j^w = 0$).

Figure 6: Effect of listing wording difference on effective starting price



Notes: The unit of analysis is the pair. The three subsamples analyzed are: 1) Pairs of auctions for new items that differ as to the inclusion of "new"; 2) Pairs of auctions that differ as to the inclusion of "disc"; and 3) Pairs of auctions that differ as to the inclusion of the disc's format. Bar heights indicate the mean effective starting price among auctions that include the word ($z_j^w = 1$) and among those that do not include the word ($z_j^w = 0$).

APPENDIX

A) Proofs

In additional appendix.

B) Description of data procedures

In the observational data, we exclude listings where the shipping fee is not provided; where more than one movie title is sold; where bidders set their identities as private, which prevents us from tracking bidding activity (chosen very rarely); when the seller delisted the auction before it ended; and auctions with a BIN option that was exercised (sellers can include a BIN option in the same auction, which disappears when the first auction bid is placed).

Based on our visual inspection of the auction titles, we used the following terms to identify a DVD as a special edition: “special,” “deluxe,” “dlux,” “gift,” “giftset,” “ltd,” “limited,” or “extended.” Similarly, we identify auctions as widescreen if the title contained the term “ws” or “wide” (“wide” picks up “widescreen”). We identify auctions as unrated if the title contains the term “unrated.”

Identifying a movie based on the eBay listing title is not straightforward. Our Java query tool was designed to be inclusive in the sense of capturing as many listings as could be reasonably expected to appear in regular users’ searches. Towards that end, we searched the title and body description of the listing for our search terms, while eBay by default only searches the listing title. Additionally, our search string included the movie title only without additional terms to narrow the search. For example, we searched for new DVDs for the 2005 movie “Batman Begins” using the string “Batman Begins” instead of “Batman Begins movie DVD (2005) new” or another permutation of possible search terms. The wording is important because of the way that eBay’s default search algorithm operates. The default algorithm is “All words, any order” such that all terms in the search string must be present and exactly as spelled in the listing for the listing to appear in search results (though are not case sensitive). This rule can generate large differences in search results based on subtle differences in search strings. For example, on September 27, 2010, a search for DVDs of "Batman Begins" using the search string "Batman Begins on DVD" returned 5 listings, while a search using "Batman Begins DVD" returned 699 listings. The reason is that most listing titles do not include the word "on." A common variation in listing titles is whether the movie release year is included. Far fewer listings are returned when year is included in the search string. For example, on September 27, 2010, "Batman Begins DVD" returned 699 listings, while "Batman Begins 2005 DVD" returned 265 listings.

There are important exceptions to the "All words, any order" rule. Generally, the algorithm returns the singular and plural versions of search terms. For example, a search for the 2008 movie "Street Kings" returns listings for the 2002 movie "The Street King" and the 2003 movie "King of the Streets." Also, for some common words such as "DVD," the search additionally returns listings that do not contain the word "DVD" but appear in a corresponding eBay listing category. For example, a search of "Batman Begins DVD" returns all listings that contain the words "Batman Begins DVD" and also listings in eBay's "DVDs & Movies" listing category that contain the words "Batman Begins" without the word "DVD."