Learning and Information Design on an Auction Platform

PRELIMINARY DRAFT

Joshua D. Higbee*

September 12, 2024

Latest draft available here

Online platforms often do not directly control users' pricing strategy, instead offering analytics and other information to help steer user behavior. I study the role of information provision by an auction platform using data from eBay auctions of children's toys. I show that as new sellers gain more experience they set lower reserve prices, earn higher revenues, and attract more bidders. I develop a model of selective platform participation where new sellers learn to set reserve prices through repeated transactions. I provide conditions under which new sellers' beliefs about the effect of reserve prices on bidder arrival are semiparametrically identified. In a dataset of eBay auctions for children's toys, estimates of the learning model suggest that new sellers underestimate the effect of high reserve prices on deterring bidder entry. Counterfactual simulations indicate that platform and seller profits improve, and more bidders enter, when the platform can shift new sellers' beliefs toward the true parameters.

*I thank Ali Hortaçsu, Dennis Carlton, Giovanni Compiani, Gunter Hitsch, Eric Richert, Brad Larsen, Avner Strulov-Shlain, Eric Budish, Chad Syverson, Devin Pope, Stephane Bonhomme, Kirill Ponomarev, and seminar participants at the University of Chicago and BYU for their helpful comments.

1 Introduction

Two-sided platforms help buyers and sellers transact by both providing a marketplace and accompanying infrastructure to successfully match users. However, these firms play a limited role in the decision problem of sellers that use their site: individual users choose which items to sell, as well as how to price and promote their listings within the platform's interface. Despite this decentralized approach to on-site transactions, many platforms provide information to sellers to help them track and optimize their business (e.g., eBay Seller Hub, Amazon Seller Central, AirBnB Smart Pricing, Walmart Seller Academy). This information directly affects sellers' strategic decisions, since sellers may be uninformed or uncertain about how to maximize profits in a new setting.

In this paper, I study an auction platform's problem of optimal information provision when sellers face uncertainty about their expected profits. I first document seller learning in a large, rich dataset of eBay auctions, implying sellers may benefit from additional information about the auction process. To study the underlying features of the platform, I combine a seller learning model with the two-sided auction platform model from Marra (2019). In this setting, sellers face a tradeoff between extracting surplus from existing bidders and attracting more bidders through lower reserve prices. I characterize the optimal reserve price of a fully informed, profit-maximizing seller and show that it generalizes the reserve price of Myerson (1981) to a setting with endogenous auction entry. When sellers are learning about bidders' entry process, however, their choice of reserve price varies with their beliefs. I provide conditions under which new sellers' beliefs are semiparametrically identified, and I estimate the auction platform model using debiased machine learning.

The question of optimal information provision is central to the platform's strategic problem. Within the marketplace, sellers face a problem similar to a standard monopoly pricing problem (Bulow and Klemperer 1996). In this context, the demand faced by each individual seller is determined by both how many bidders enter each auction and how much those who enter value the item being sold. Sellers must therefore know both the bidder entry process and the distribution of bidder valuations to maximize profits from the items they list. If sellers have incomplete information about their demand curve, the platform may wish to correct sellers' beliefs and benefit from their more informed choices. However, the platform may benefit from information asymmetries among its users just as it may benefit from setting different fees on both sides of the market (Rochet and Tirole 2003). Thus, sellers' beliefs and learning about the arrival process–and the platform's role in influencing these beliefs–shape outcomes for sellers, bidders, and the platform itself.

This analysis is motivated by empirical patterns in a large dataset of eBay auctions for children's toys from before the introduction of eBay's Seller Hub. I demonstrate that new sellers set higher reserve prices and earn lower revenues than their more experienced counterparts. There is evidence that is suggestive of both selection and learning among new sellers. First, new sellers that remain on the platform set lower prices and earn higher revenues in their first auctions than new sellers that list a few items and then exit. Restricting attention to sellers that remain on the platform reveals the same pattern: new sellers lower their reserve prices and increase revenues as they gain more experience.

Reduced-form evidence alone, however, cannot disentangle how much of sellers' pricing behavior is attributable to selection and how much is due to seller learning. This is because both sellers' beliefs and their private valuations for an item affect their choice of reserve price. By explicitly modeling the arrival and decision process of both bidders and sellers, I characterize key features of the auction platform that drive seller behavior. In particular, high reserve prices reduce the expected benefit to bidders of entering auctions, which leads to lower entry. I show that sellers can overcome this by setting a reserve price that balances the standard "surplus extraction" motive of Myerson (1981) against the "bidder attraction" motive inherent to the platform game. This provides theoretical justification for the lower reserve prices set by more experienced sellers. Further, bias in seller beliefs about bidders' entry process may cause sellers to incorrectly conclude that it is unprofitable to list certain items for auction. I model how beliefs affect seller behavior on both the intensive margin (through reserve prices) and on the extensive margin (through selective entry).

In order to quantify the effect of information on seller actions, I show seller beliefs about bidders' entry process are semiparametrically identified from choice data within a single period. I highlight similarities between the identification of individual beliefs and the identification of random coefficients models (Fox et al. 2012), and show how variation in sellers' choice of reserve price traces out sellers' beliefs about their profit function. I also propose a likelihood approach to estimate bidder values, their arrival process, and heterogeneity in item values. This approach solves two key challenges in my setting. First, since these are ascending auctions, the distribution of observed bids is higher (in a first-order stochastic dominance sense) than the true bidder value distribution. I address this by explicitly modeling the arrival process and its influence on the distribution of the highest observed bids. Second, each item is characterized by its auction title, which is written by the seller. Since the set of possible item titles is extremely high dimensional, I embed deep neural networks within the likelihood model to flexibly estimate item-level heterogeneity along with high-level parameters such as the bidder arrival process. However, this approach can lead to significant bias in the estimated parameters of interest due to over-fitting the model to the data. In line with the literature on debiased machine learning (Farrell, Liang, and Misra 2020; Chernozhukov et al. 2022; Ichimura and Newey 2022), I derive a Neyman-orthogonal score and use it to estimate the auction model. Thus, I obtain asymptotically consistent estimates of the model parameters despite the large dimension of the dataset.

My estimates illustrate the importance of several underlying features of the model. Bidders face a non-trivial time cost of entering each auction and inspecting the listing. Consistent with the model, this means higher reserve prices have a strong negative effect on the expected number of bidders. This effect is distinct from the mechanical effect of increasing the public reserve price, which excludes potential bidders from submitting a bid even if they enter the auction. New sellers underestimate the bidder-deterrence effect, but among those that continue to list items for auction, sellers soon set prices that are broadly consistent with accurate beliefs about the bidder arrival process.

These results have important strategic implications for the platform and its decision of how much information to provide. The platform can choose to provide new sellers with data from past auctions, from which they can learn about the true bidder entry process. At the same time, the platform chooses what percentage of revenue it will charge from both new and experienced sellers, who are already familiar with the bidder arrival process and set prices optimally. My estimates quantify how much and in what direction new sellers will adjust their beliefs about the bidder entry process in response to new information. When provided with a random sample of other auctions on the platform, the average new seller is able to update their beliefs toward the true parameters without incurring the time and monetary cost of starting to use the platform. This allows the platform to change its fee structure to increase platform profits from the baseline, though it may be optimal for the platform to further increase profits by exploiting new sellers' information gap.

Empirically estimating new sellers' beliefs about the bidder entry process is critical, as a platform's optimal information structure may be ambiguous without empirical evidence. First, while platforms may have increasing economies of scale in analyzing data, it is still costly to provide users with additional information. This may be exacerbated by users' low willingness to pay for additional information: sellers who do not anticipate changing their beliefs will see little value in purchasing additional data. Thus, despite the cost, platforms may prefer to offer such information for little to no fee, if they offer it at all. Second, it is possible for the platform to help one side of the market without harming the other: correcting sellers' beliefs yields improved profits for sellers while inducing more bidder entry. The effects of this may be magnified by the two-sidedness of online platforms: increasing the surplus of one side of the market may lead to more entry on both sides. Finally, since the users' decision problem may change with new information, the optimal fee structure on a two-sided platform (as studied under perfect information in e.g. Rochet and Tirole 2003 and Klein et al. 2005) may also change.

1.1 Related literature

This paper contributes to the large and growing literature on agent learning. Much of this literature focuses on learning-by-doing and the extent to which more experienced agents are able to leverage information to improve outcomes (Simonsohn 2010; Haggag, McManus, and Paci 2017; Strulov-Shlain 2021; Tadelis et al. 2023). In particular, Huang, Ellickson, and Lovett (2020) shows how firms entering into a new market adjust to market signals and set prices accordingly. I document similar behavior among new sellers on an auction platform and contrast this behavior to that of more experienced agents. The pattern of new sellers lowering reserve prices over time runs opposite to the mechanism in Foster, Haltiwanger, and Syverson (2016), where new firms temporarily set lower prices to increase demand in future periods, and it persists when controlling for seller reputation variables that are observed by bidders. Other work addresses the theory of optimal behavior under uncertainty, especially in sequential games with updating (Rothschild 1974; Keller and Rady 1999; Hitsch 2006). I contribute to the literature on estimating agent beliefs and learning process, as in Erdem and Keane (1996) and Kim (2020); I also provide semiparametric identification results for beliefs under learning as in Lu (2019), but without requiring that the learning process be Bayesian.

I also address the field of auction design and optimal pricing in auctions. Auctions are a particularly well-suited application for information design in platform markets due to the rich auction theory literature that explores bidders' and sellers' optimal strategies in a variety of settings. Additionally, a variety of empirical tools facilitate an empirical analysis with which to test and quantify theoretical results. In particular, I adapt the endogenous auction platform model of Marra (2019), which relates to other models of auctions with endogenous entry such as Levin and Smith (1996), to a setting with seller uncertainty. My model also develops an insight from Engelbrecht-Wiggans (1987) (that optimal reserve prices should account for their effect on bidder arrival) into a new reserve price condition that nests that of Myerson (1981); this reserve price also shows the tradeoff between the benefit of attracting another bidder (as in Bulow and Klemperer 1996) and expected surplus extraction when arrival is both endogenous and stochastic. Existing literature further shows that auction participants may not always behave according to theory, even in laboratory settings or when they are large, sophisticated firms (Davis, Katok, and Kwasnica 2011; Ostrovsky and Schwarz 2016); my application examines small firms in a real marketplace.

Finally, this paper relates to the growing literature on two-sided markets. Existing work has studied the question of cross-subsidization in optimal platform design both theoretically and empirically (Rochet and Tirole 2003; Klein et al. 2005; Gomes 2014; Jullien, Pavan, and Rysman 2021; Marra 2019). I consider the problem of optimal platform fees jointly with the problem of information design to explore how a platform's ability to shape users' information affects its own profits as well as participants' welfare. Other work also examines the role of information and learning on platforms; Mela, Roos, and Sousa (2023) documents firms learning to advertise on a platform after initially overestimating the effectiveness of advertising, and Foroughifar (2023) studies AirBnB hosts' beliefs about and use of smart pricing tools. More broadly, studies of search, ranking, and recommendation systems consider the same problem of influencing user behavior through non-price mechanisms, though these generally focus on the buyer side of a two-sided platform (Bronnenberg, Kim, and Mela 2016; Compiani et al. 2022; Xu, Deng, and Mela 2022; Hodgson and Lewis 2023).

2 Setting and descriptive evidence

The data for this paper comes from a sample of eBay auctions used in Resnick and Zeckhauser (2002), and spans from January to June 1999. eBay is well known as a platform for users to buy and sell items; at the time of the data, auctions were the only mechanism used on the site. This data also preceeds the introduction of eBay's data analytics service "Seller Hub" in 2016.

Before examining the data in more detail, I present a simplified outline of the eBay auction process. Sellers choose to list an item for auction, and choose a starting minimum bid and (if desired) a secret reserve price along with an item description. Prospective bidders can find listed items on a search page, along with some information about the current price and the number of bids submitted, and then choose to click into the item page. Bidders may then observe the current minimum bid along with seller information and an indicator for whether the secret reserve price (if any exists) has been met. Bidders submit their bids to eBay, which proceeds as a second-price ascending auction (where the current minimum bid is the second-highest of existing bids and the initial minimum bid). Examples of the search and item pages are presented in Figure 1.

2.1 Data

Due to the prevalence of antique and custom items within the full dataset, I restrict attention to one of the more popular categories: a brand of stuffed animals called Beanie Babies (BBs). There are approximately 1 million BB auctions in the dataset, corresponding to about 2.7 million bids. All distinct varieties of BBs were produced by a single company, which ensures some level of homogeneity among the listed items. Table A1.1 presents various summary statistics for the sample of items in the analysis dataset.

Several features make this dataset attractive for empirical analysis. First, both the highest

Figure 1: Example of eBay search and item pages

Notes: This figure was retrieved from the Wayback machine, and has been edited to conserve space. The search and item pages are from 2001 and 2005, respectively. The sellers' ability to feature their item can be seen here - featured items are listed at the top of the page, while other items are highlighted and/or have bold titles. The seller's total feedback score and positive feedback rating are visible, as is the current bid, starting bid, and and indicator for the reserve not being met by the current bid.

bids and secret reserve prices are recorded for each item in the dataset. Online auction datasets frequently impute secret reserve prices from observable indicators such as the "reserve not met" sign in Figure 1; since my focus is on sellers' choice of reserve price, it is helpful to obtain accurate measurements of this choice variable. Highest bids are similarly unobserved in many studies of online auctions due to the ascending minimum bid only depending on the second-highest bid. As will be shown later, the first-highest bid being observable is helpful for identifying and estimating seller beliefs. Other data such as the time of the auction, any promotional choices made by the seller, and seller-provided item descriptions are also included in the dataset.

Additionally, supplemental data on users' feedback reveals all positive, negative, and neutral ratings for accounts. Importantly, this feedback history dates back to the beginning of eBay in 1995; this allows me to construct the reputation variables that are visible to prospective bidders, specifically feedback scores (defined as the number of all feedback events) and rating (defined as the percent of feedback events which are positive). Unless otherwise noted, I use the inverse hyperbolic sine of feedback scores to address the significant skew in user feedback. While feedback scores are not a perfect measurement of the number of auctions in which sellers have participated, they are often used as a proxy for user experience on eBay (see e.g. Simonsohn 2010).

2.2 Trends among new sellers

To motivate the model, I highlight trends among new sellers' choices and outcomes. I define new sellers as all accounts who have no recorded feedback before the start of the dataset and who list at least one item for sale. Similar to Kim (2020), I use a reduced-form test of the differences between new and experienced sellers (those with above-median feedback scores at the start of the data). In particular, for each seller ℓ and auction k, I estimate

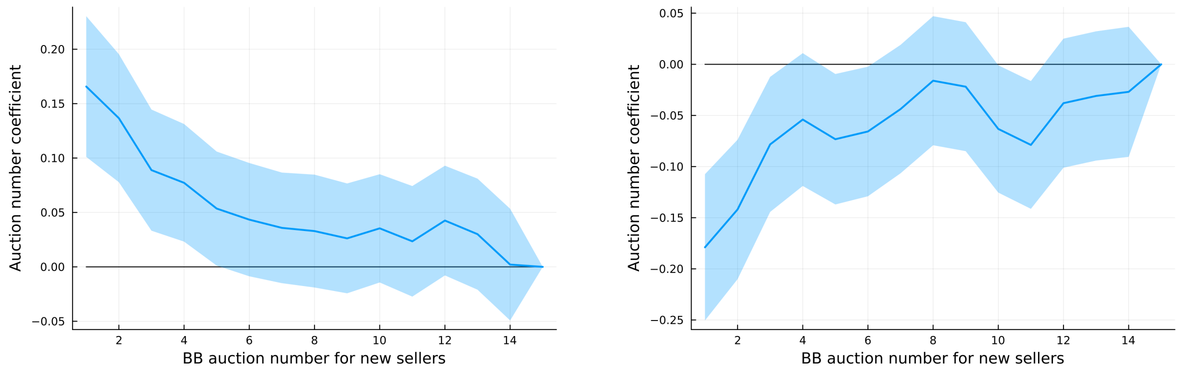
$$y_{\ell k} = \sum_{k=1}^{K-1} \alpha_k \cdot \mathbb{1}[\ell \text{ is new}] + x'_{\ell k}\beta + \lambda_\ell + \eta_{t(k)} + \epsilon_{\ell k}$$

where $y_{\ell k}$ is a variable of interest, $x_{\ell k}$ contains predicted item values, seller feedback, and seller ratings, and λ_ℓ and $\eta_{t(k)}$ are respectively seller and month fixed effects. The variables of interest are seller revenue net of fees and the *effective reserve price*, defined as the maximum of the secret reserve price (if one exists) and the starting minimum bid. Throughout, I will homogenize variables such as reserve prices and revenues by dividing the raw number by the predicted item value. The process for estimating item values is detailed in section 4.1; additional plots in Appendix A1 replicate these trends without depending on the predicted values.

Figure 2 plots regression coefficients α_k when restricting the sample of new sellers to those with at least $K = 15$ auctions in the data.[1] These new sellers initially set higher prices and earn lower revenues than they do in later auctions. This pattern is consistent with seller learning when initial beliefs are biased toward higher prices. Further, by examining only those who remain on the platform for at least 15 auctions and controlling for persistent seller heterogeneity, these trends do not simply reflect early exit by high-value sellers.

[1] I chose 15 auctions to avoid including sellers who may have few items in their possession and no interest in long-term trading, as well as to not have too short a panel for estimating seller fixed effects. The trends are similar for different values of K.

Figure 2: Regression coefficients α_k of auction experience on variables of interest

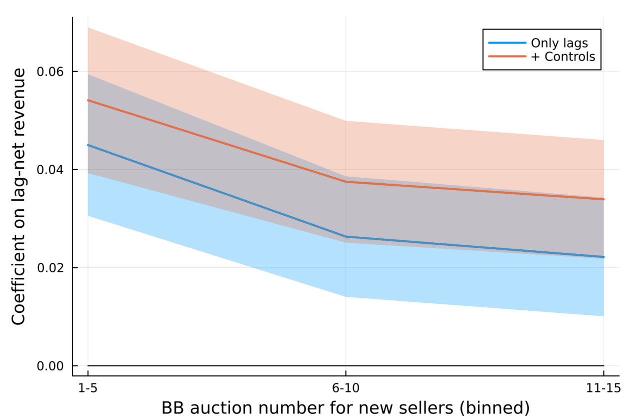


(a) Effective reserve (homogenized) (b) Net revenue (homogenized)

Notes: These regressions pool 1,639 new sellers' first 15 auctions with all auctions by 5,165 experienced sellers (defined as those with ≥47 auctions at the start of the data, which is the 75th percentile of initial experience). The sample is limited to sellers with at least 15 auctions in the data. The results are similar when using different values of K.

Standard learning models also predict that the value of additional information decreases as sellers obtain more experience, and that seller beliefs converge toward toward the true parameter. To examine variability in seller choices over time, I regress current-auction reserve prices on lagged net revenue and lagged reserve prices (again conditioning on sellers with at least 15 auctions in the data). Figure 3 plots the coefficients of lagged net revenue from this regression, binned by the auction number among new sellers. New sellers' current prices are correlated with past revenue signals, and the magnitude of the lag coefficient is larger in early auctions. This is consistent with subsequent auctions containing relatively less information for more experienced sellers.

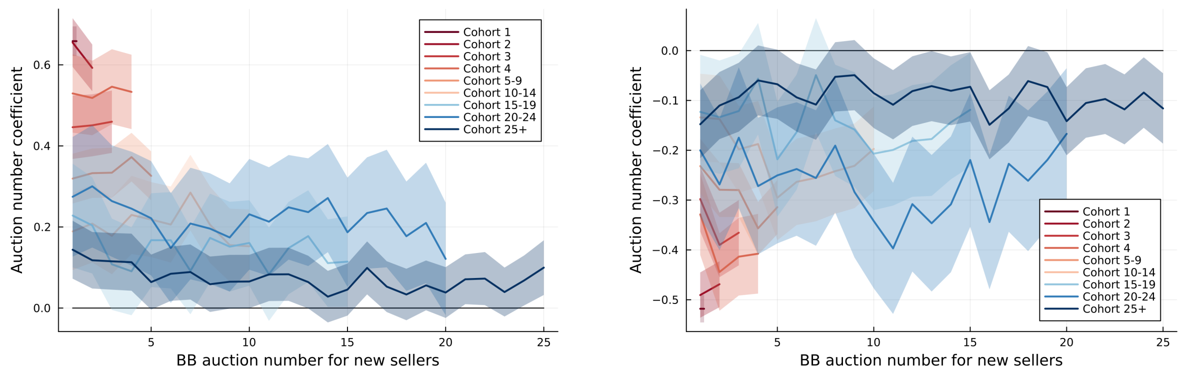
Figure 3: Coefficients from regressing prices on lagged revenues



Notes: These are the coefficients when regressing current-period effective reserve price on lagged revenue, multiplied by indicator functions for new sellers being in the first 1-5, 6-10, and 11-15 auctions in the data (among new sellers with at least 15 auctions in the data and experienced sellers with >75th percentile of experience at the start of the data). The controls include month fixed effects, feedback percentage, and predicted item value.

Similar trends hold when examining new sellers more broadly, though the estimates are far noisier. To examine the effect of selection on new seller outcomes, I separate new sellers into cohorts based on the number of items they list. In Figure 4, I plot the average difference between new sellers' and experienced sellers' reserve prices and net revenue for each of new bidders' first k auctions. Consistent with selection, new sellers who list relatively few items are also those with significantly higher reserve prices and lower revenues.

Figure 4: New seller trends by cohorts of the number of items listed (difference from experienced seller averages)



(a) Effective reserve (homogenized) (b) Net revenue (homogenized)

Notes: These figures represent a simple difference in means between new sellers in each cohort and experienced sellers (defined as sellers with >75th percentile of experience at the start of the data). This does not include any controls other than month fixed effects.

Additional results in Appendix A1 highlight other patterns in seller behavior over time. Sellers exhibit some similar trends in non-price choice variables, notably whether to feature an item and when and for how long to list it. However, the magnitude of the trends are generally small, which motivates sellers' pricing decision as the choice variable. I also show that, while there are some trends in item descriptions, some of the most prominent focus on the item reserve price (in particular, noting the lack of a secret reserve price).

2.3 Alternative explanations

Before moving to the auction model, I consider other mechanisms that might drive the trends shown above. First, sellers might strategically choose the order of their listings when starting their account, perhaps starting with items they value more highly. However, this strategy runs counter to any dynamic considerations like those discussed in Foster, Haltiwanger, and Syverson (2016).

Selling low-price items upon entering the platform would yield higher sale probabilities, allowing sellers to increase their reputation scores and potentially earn more in subsequent auctions.

Sellers may face a different set of bidders in their initial auctions, among whom it may be optimal to set higher prices than for later auctions. It is difficult to directly analyze the distribution of bidder values in the absence of a model, since the increasing minimum bid and endogenous entry of bidders create a selection problem. To overcome this, I restrict attention to the first and second highest bids (where they exist) and flexibly control for the number of observed bids with fixed effects. Figure A1.4(a) shows generally flat trends in the first and second highest bids in new sellers' auctions. This suggests that higher reserve prices are not driven by differentially higher bids by sellers.

Finally, sellers may have systematically higher values for all items upon entering the platform. In this dataset, I can observe sellers whenever they bid for other items and test whether their bids change as they gain more experience on the platform. Figure A1.4(b) shows the bids of new sellers for other listings, and illustrates that new bidders do not place higher (or lower) bids for other items upon entering the platform. This is consistent with the underlying distribution of seller valuations remaining constant throughout new sellers' first auctions on the platform.

3 Auction platform model

The empirical patterns in the data may be driven by many factors, including both selective seller entry and seller learning. Further, any seller learning must be about some feature of their environment. To directly study the seller problem and understand their learning process, I develop an auction platform model that allows for seller learning. The model builds on the two-sided endogenous entry model from Marra (2019) by allowing seller actions to vary with their individual beliefs about their profit function, and will note where the models differ significantly. I begin by introducting the following notation and assumptions.

A large number $\mathbf{N}_S$ of sellers and $\mathbf{N}_B$ of bidders have access to a monopoly auction platform and can choose to participate. Each potential bidder i has valuation v_{ij} for each item j that may be traded on the platform, where $v_{ij} \sim F_B$. Potential sellers ℓ have outside option values $v_{0\ell j}$ for each item j they possess, where $v_{0\ell j} \sim F_S$. Each item has auction-level observables X_j. I assume all prospective sellers and bidders know the valuation distributions F_B and F_S, which satisfy the following assumption.

Assumption 1. The value distributions F_B and F_S are absolutely continuous and have connected support. Bidder values v_{ij} are independent from $v_{i'j}$ for all $i \neq i' \in \{1, ..., \mathbf{N}_B\}$, and seller values $v_{0\ell j}$ are independent from v_{ij} for all $i \in \{1, ..., \mathbf{N}_B\}$ and $\ell \in \{1, ..., \mathbf{N}_S\}$.[2] Further, the bidder value distribution F_B satisfies the strict monotone hazard rate property (i.e., $\frac{f_B(x)}{1-F_B(x)}$ is strictly increasing in x on the support of F_B). Finally, dependence on item j's characteristics takes the form $v \cdot \exp(\gamma(X_j))$, where X_j is exogenous to all values v, which are drawn from the players' respective distributions (F_B or F_S).

Throughout the discussion of the model, I focus on homogenized item values, which is equivalent to setting $\gamma(X_{jt}) = 0$ for all items. Since each item j is sold by a single seller ℓ, I omit dependence of values and other terms on j and ℓ where possible.

I assume all sellers know their values v_0 for the item they own. Sellers can choose to list the item for auction after incurring an item-specific entry cost $c_S^E \overset{\text{i.i.d.}}{\sim} F_{c_S^E}$, which is also independent from sellers' private values v_0. After deciding to list the item, sellers also choose an effective reserve price r and minimum bid m. For the purposes of the model, I assume m is exogenously drawn between 0 and r, and unlike Marra (2019), I assume that r is observable to all prospective bidders.[3] Potential bidders can see the item on a listing page and decide whether or not to enter the auction; if they do so, they incur entry cost c_B^E and only then learn their value v_i for the item. Bidders may then costlessly submit a bid. The bidder with the highest bid wins the item if their bid exceeds the reserve price, and the transaction price p is equal to the highest of the effective reserve price and the second highest bid.

[2]Other work explores the relationship between reserve prices and outcomes in private and common values settings (Quint 2017).

[3]In this dataset, only 23.9% of auctions have a secret reserve price, while 89% have an minimum bid higher than the eBay default of \$1. Thus, this simplifying assumption is reasonable in the present setting. The increasing minimum bid rule on eBay means the "reserve not met" indicator disappears when at least one bid passes the reserve price, at which point the minimum bid has often increased from its starting value. Further, Katkar and Reiley (2007) documents that the reserve price affects bidder entry even when it is still secret through the "reserve not met" indicator. Higher reserve prices are less likely to be met by future bidders, so on average a higher secret reserve price will be flagged for a longer period. A broader literature relates to the choice between public and secret reserve prices (Hasker and Sickles 2010).

Bidder and seller behavior is also affected by the cost structure of using the platform, denoted by a vector c of all associated costs and fees. In particular, entry costs c_B^E and c_S^E for both bidders and sellers are decomposed into time costs c_B^T and c_S^T and insertion fees c_B^I and c_S^I, respectively. The insertion fees are paid directly to the platform when a seller lists an item or when a bidder enters an auction, regardless of whether the item is sold. The platform can also impose bidder and seller fees c_B^P and c_S^P. If the item is sold, the highest bidder pays $(1 + c_B^P)p$ and the seller receives $(1 - c_S^P)p$, so the platform also receives revenue $(c_B^P + c_S^P)p$ from each successful sale.

The rest of the section is divided into three parts. I first consider bidder strategies, conditioning on seller behavior, since a profit-maximizing seller considers the behavior of bidders that may enter their auction. I then examine the seller strategies and how they depend on sellers' beliefs about bidder behavior. Finally, I review the conditions on both bidder and seller behavior that must hold simultaneously in equilibrium.

3.1 Bidder strategies

I first focus on the the bidding strategy of actual bidders who enter the auction. This implies a continuation value of entering in an auction, which pins down bidders' optimal decision when deciding whether to enter an auction.

3.1.i Bidding stage

All $\tilde{N}$ bidders who have entered the auction face no cost to submitting their bid, but may be constrained from doing so by the minimum bid m. Following Vickrey (1961) and Marra (2019), all bidders with value v_i will submit bids $\frac{v_i}{1+c_B^P}$ as long as their bid exceeds the current value of m. Any bidder with $v_i < (1 + c_B^P)m$ will not bid at all. While this poses a selection problem for estimation in section 4, it has no effect on the outcome of the auction game.

I abstract from any potential learning and uncertainty in bidder strategies. This is because eBay provides an automatic bidding tool that increments the minimum bid up to the maximum value a bidder reveals they are willing to pay. Thus, eBay already implements the optimal bidding rule for all bidders via algorithm. This also allows for more tractable modeling of bidder behavior, both for the researcher and for the sellers' mental model of bidder behavior.

3.1.ii Entry stage

In this setting, in contrast to Marra (2019), the reserve price is public. This means that even though bidders do not know their value before entering the auction, they form expectations about their expected surplus from entering the auction based on the price they see at the search stage. More formally, define the fee-adjusted reserve price faced by bidders as $r^B \equiv (1 + c_B^P)r^*$ and let Λ parameterize bidder arrival. Then a potential bidder's *ex ante* expected surplus from entering an auction is

$$\pi_B(r \mid \Lambda, c) = \sum_{n=1}^{\mathbf{N}_B - 1} \underbrace{\frac{1}{n}}_{\text{(i)}} \cdot \underbrace{\mathbb{E}\Big[v_{n:n} - (1 + c_B^P)\max\{v_{(n-1):n}, r^*\} \;\Big|\; v_{n:n} \geq r^B\Big]}_{\text{(ii)}} \cdot \underbrace{(1 - F_B(r^B)^n)}_{\text{(iii)}} \cdot \underbrace{\mathbb{P}[\tilde{N} = n \mid \Lambda]}_{\text{(iv)}} \tag{1}$$

where $v_{\ell:n}$ is the ℓth highest out of n realizations of v_i. The four components of π_B are (i) the probability that any given bidder has the highest value, (ii) the expected surplus when the highest bidder wins, (iii) the probability that the highest bid exceeds the fee-adjusted reserve price, and (iv) the probability that n bidders enter at the auction.

The following proposition characterizes the equilibrium of the bidder entry game. The existence and uniqueness of equilibrium in this setting follows from Marra (2019), but the presence of a public reserve price yields a new result: the expected number of bidders is decreasing in the posted reserve price.

Proposition 1. Assume bidder arrival is i.i.d. Poisson with mean $\Lambda(r)$. Then the bidder entry equilibrium exists and is unique. Further, $\frac{\partial \Lambda}{\partial r} < 0$. (Proof in Appendix A2)

The intuition for this result is as follows. Potential bidders have zero expected profit (net of time cost) from entering an auction, i.e.

$$0 = \pi_B(r \mid \Lambda, c) - c_B^E \tag{2}$$

since any positive profit will induce additional entry and negative profit will cause excess bidders to leave. The expected surplus in an any auction is declining in the number of competing bidders, since additional bidders both increase the expected price paid and lower the probability that a new bidder will win the item. Expected surplus is decreasing in the equilibrium number of expected bidders, Λ, and properties of the Poisson distribution imply that a unique value of Λ satisfies the expected zero profit condition for each r. Finally, expected bidder surplus declines with r, so fewer bidders enter when they expect more competition from the seller.

For tractability, I parameterize the mean number of bidders Λ in equilibrium as a function of some vector δ_0. Proposition 1 shows that for each r there is a different expected number of bidders that enter the auction. Without loss of generality, I write the expected number of bidders as

$$\Lambda(r \mid \delta_0) = \exp(\delta_{0,1} + \delta_{0,2}\rho(r)) \tag{3}$$

for some strictly increasing function ρ that is determined by the zero-profit condition (2). Importantly, Proposition 1 implies $\delta_{0,1} < 0$: any increase in the reserve price will reduce the expected number of bidders that enter the auction. Combined with the assumption that bidder arrival is Poisson, the probability that n bidders enter the auction is

$$p_n(r \mid \delta_0) = \frac{\Lambda(r \mid \delta_0)^n \exp(-\Lambda(r \mid \delta_0))}{n!} \tag{4}$$

for $n = 0, 1, 2, ...$ given $\Lambda(r \mid \delta_0)$.

3.2 Seller strategies

I first show how the platform's pricing structure and endogenous bidder entry shape sellers' choice of optimal reserve price. I then extend this result to allow for seller uncertainty and learning, and conclude with sellers' entry problem. Throughout, I treat the sellers' item-specific entry cost c_S^E as fixed for a given auction.

3.2.i Platform reserve price under perfect information

I begin with a general form of the seller profit function to fix ideas. Sellers' expected profit, conditional on δ_0, is given by

$$\Pi(v_0, r \mid \delta_0, c) = (1 - c_S^P) \cdot \underbrace{R(r \mid \delta_0)}_{\mathbb{E}[\text{Revenue}|r,\delta_0]} + v_0 \cdot \underbrace{K(r \mid \delta_0)}_{\mathbb{P}[\text{Keep}|r,\delta_0]} - c_S^E$$

where the precise functional forms of R and K follow from the second-price auction literature and the Poisson arrival function.[4] Taking first-order conditions, the optimal interior reserve price r^* satisfies

$$\psi(r^* \mid \delta_0, c) \equiv \frac{-(1 - c_S^P) R_r(r^* \mid \delta_0)}{K_r(r^* \mid \delta_0)} = v_0$$

where ψ is the virtual type function mapping bids to the space of seller values. Applying the Poisson assumption and the relevant functional forms, and then rearranging the terms of R_r and K_r, yields the following result:

Proposition 2. Assume bidder arrival is Poisson with mean $\exp(\delta_{0,1} + \delta_{0,2}\rho(r))$. Then (recalling that the fee-adjusted reserve price faced by bidders is $r^B \equiv (1 + c_B^P) r^*$) the optimal interior reserve price satisfies

$$\frac{v_0}{1 - c_S^P} = \left[r - \frac{1 - F_B(r^B)}{(1 + c_B^P) \cdot f_B(r^B)} - \frac{W_R(r \mid \delta_0)}{f_{\max}^B(r \mid \delta_0)} \right] \frac{f_{\max}^B(r \mid \delta_0)}{f_{\max}^B(r \mid \delta_0) + W_K(r \mid \delta_0)} \tag{5}$$

where $f_{\max}^B(r \mid \delta_0) \equiv (1 + c_B^P) \cdot \sum_{n=1}^{\bar{N}_B} p_n(r; \delta_0) F_B(r^B)^{n-1} f_B(r^B) n$ is the scaled density of the highest bid at r given δ_0, and $W_R(r; \delta_0) \equiv \sum_{n=0}^{\mathbf{N}_B} (\frac{\partial p_n(r;\delta_0)}{\partial r}) R_n(r)$ and $W_K(r; \delta_0) \equiv \sum_{n=0}^{\mathbf{N}_B} (\frac{\partial p_n(r;\delta_0)}{\partial r}) K_n(r)$ are weighted averages of the expected revenue and keep probabilities for each n.

[4] Formally, these functions are written $R(r \mid \delta_0) \equiv \sum_{n=0}^{\mathbf{N}_B} p_n(r \mid \delta_0) R_n(r)$ and $K(r \mid \delta_0) \equiv \sum_{n=0}^{\mathbf{N}_B} p_n(r \mid \delta_0) K_n(r)$, where $r^B = (1 + c_B^P) r$ and

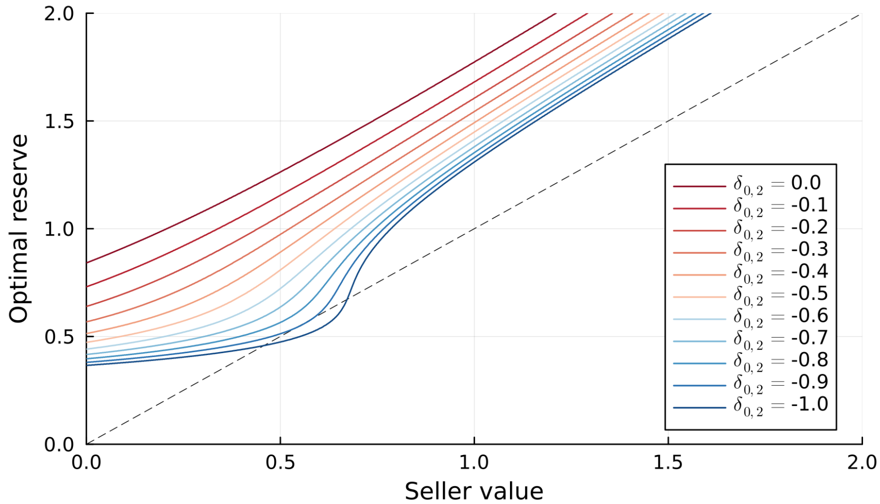
$$R_n(r) \equiv nr(1 - F_B(r^B)) F_B(r^B)^{n-1} + \frac{n(n-1)}{1 + c_B^P} \int_{r^B}^{\infty} z(1 - F_B(z)) F_B(z)^{n-2} f_B(z) dz$$
$$K_n(r) \equiv F_B(r^B)^n$$

Note that when $\delta_{0,2} = 0$, the arrival process does not depend on r and $W_R = W_K = 0$. When the bidder and seller fees are also zero ($c_S^P = c_B^P = 0$), equation (5) reduces to the Myerson (1981) optimal reserve price formula.

To provide intuition for this pricing rule, Figure 5 plots the implied reserve price from Proposition 2 for different values of the bidder arrival coefficient $\delta_{0,2}$. The figure starts from the baseline case of $\delta_{0,2} = 0$ and shows increasingly negative values of $\delta_{0,2}$ in comparison. As $\delta_{0,2}$ decreases, sellers should optimally lower the markup in their reserve price. In this particular case, it may even be optimal to set a reserve price lower than the seller's own value if the bidder deterrence effect is sufficiently strong. The precise shape of the optimal reserve price function depends on the intercept $\delta_{0,1}$ of the log-mean of average bidders, as well as the shape of the bidder value distribution F_B and the function ρ that is determined by bidders' zero profit condition (2).

This equation generalizes several results from the related auction literature. As previously noted, it nests the Myerson (1981) reserve price formula for $\delta_{0,2}$ and carries with it the same intuition: sellers may set a higher reserve to extract additional surplus from bidders. Sellers' influence over the arrival process means they explicitly weigh the expected benefit of surplus extraction against the expected benefit of a potential additional bidder (though an extra bidder would be better in expectation, as in Bulow and Klemperer (1996), this arrival is not guaranteed). Thus, this equation represents a version of the argument in Engelbrecht-Wiggans (1987), in that sellers may wish to lower the reserve to attract more bidders. Since there is some probability that fewer bidders arrive, however, the seller may still "protect" some expected surplus by setting a non-trivial reserve price.

Figure 5: Optimal reserve price for varying reserve price coefficients $\delta_{0,2}$



Notes: These figures show optimal reserve price functions for varying parameters of $\delta_{0,2}$, keeping $\delta_{0,1} = 1.0$ fixed, where log-bidder values are normally distributed with mean 0 and variance 0.4, and $\rho(r) = r \cdot F_B(r)$. The dashed 45-degree line represents all values where the seller value is equal to the optimal reserve price.

3.2.ii Platform reserve price under uncertainty

I now assume sellers may not know the true parameter δ_0. This uncertainty arises because δ_0 is a part of bidders' equilibrium play in the "search" rather than "item" stage of the game. Any uncertainty about or inattention to the platform's search algorithm, bidders' search strategies, or bidders' search costs could therefore contribute to uncertainty about sellers' own effect on prospective bidders' search process. For example, Simonsohn (2010) documents that eBay sellers may not understand the impact of competition on their own profits and over-enter when market activity is high. This competition neglect is related to seller behavior in this setting, as sellers may set reserve prices too aggressively (in essence, competing) for their own items. These sellers may fail to realize the extent to which this behavior crowds out potential bidders, each of whom may choose another way to spend their time instead of entering an auction with an uncertain payoff.

The previous derivations can be extended straightforwardly in the case where sellers have some belief density b about the true value of δ_0. In an abuse of notation, we define subjective expected profit as $\Pi(v_0, r \mid b, c) \equiv \int \Pi(v_0, r \mid \delta, c) b(\delta) d\delta$. The respective profit function components $R(r \mid b)$ and $K(r \mid b)$ are defined similarly, implying the subjective virtual type function $\psi(\cdot \mid b, c)$ from rearranging the first-order condition of $\Pi(v_0, r \mid b, c)$ with respect to r. For tractability, I assume all sellers are myopic, so they only maximize current-period profits conditional on their beliefs b

and do not actively experiment.[5]

As is standard in models of learning, sellers also have a model of the true data-generating process, from which they learn about the unknown parameter δ_0. I assume sellers update their beliefs after every item they list for auction. All sellers believe profit draws y are generated by the process

$$y = \Pi(v_0, r \mid \delta_0, c) + \epsilon \tag{6}$$

where ϵ is drawn i.i.d. from some distribution $F_{\epsilon|r,v_0}$ that is known to sellers. Each auction, sellers observe the associated data $\mathbf{D} = \{y, v_0, r\}$, and have the likelihood $l_S(\delta \mid \mathbf{D})$ as implied by the noise distribution $F_{\epsilon|r,v_0}$. After each auction, sellers use a deterministic transition function $\mathcal{T}$ to update their prior beliefs b to the posterior b':

$$b'(\delta) = \mathcal{T}(b(\delta), \mathbf{D}) \tag{7}$$

Thus, the evolution of sellers' reserve price strategies depends entirely on their beliefs b as updated after each auction, which in turn are driven by variation in data observed after each auction. I note that sellers may be Bayesian, though this is not necessary for the conclusions of the model.

3.2.iii Entry stage

Prospective sellers will enter the platform as long as their net expected surplus (over keeping the item) is positive. This yields the following inequality for sellers with beliefs b:

$$\Pi(v_0, r^* \mid b, c) - v_0 \geq 0 \tag{8}$$

I first show conditions under which r^* is increasing in v_0. This is common and easily verified under the functional forms assumed in much of the related literature; I make this explicit because it

[5] This assumption is used by Huang, Ellickson, and Lovett (2020) to simplify the analysis of firms learning from demand signals. This also relates to the "anticipated utility model" discussed in Cogley, Colacito, and Sargent (2007); in their setting, the model without active experimentation is a good approximation for the fully dynamic Bayesian model.

depends on the underlying beliefs b and is important for subsequent propositions.

Proposition 3. Assume $K_r(\cdot \mid b) > 0$. Then r^* is increasing in v_0 and the virtual type function $\psi(\cdot \mid b, c)$ is increasing. (Proof in Appendix A2)

Given the functional form of R and the monotonicity of r^* in v_0, the gains from trade $\Pi(v_0, r^* \mid b, c) - v_0$ are strictly decreasing in v_0. This implies a threshold rule for sellers, such that sellers with a private value above that threshold will not list an item for auction. Each seller's entry threshold does not depend on that of other sellers since each auction's reserve price is public knowledge, and potential bidders enter up to their expected zero profit condition. This yields the following characterization of the seller entry problem.

Proposition 4. There exists a unique threshold $\bar{v}(b, c)$ for each belief density b such that sellers with beliefs b will list their item for auction if and only if $v_0 \leq \bar{v}(b, c)$.

That is, sellers will only select into the platform if their values are sufficiently low given their beliefs. This holds regardless of heterogeneity across seller beliefs, and heterogeneity in beliefs does not directly impact the bidder arrival problem: bidders are only impacted by the reserve price r^* instead of the underlying values of v_0 and b.

This threshold condition is similar to other results in the literature, and corresponds to Marra (2019) when beliefs b are common across all sellers and a point mass on the true parameters. I weaken the assumption that this threshold is objectively correct under the true item values: it needs only be optimal according to each seller's private beliefs about the true arrival process. This ensures that a unique equilibrium exists for the seller entry game even under heterogeneous and potentially biased beliefs among sellers.

3.3 Equilibrium definition

Before proceeding, I review the conditions for both bidders and sellers that must hold in equilibrium. Given a distribution F_B of bidder values, a distribution F_S of seller outside option values, a fixed bidder entry cost c_B^E, a distribution $F_{c_S^E}$ of seller entry costs, bidder and seller transaction fees c_B^P and c_S^P, initial ($t = 0$) prior beliefs b_0 about the parameters of bidders' entry process, and an updating rule $\mathcal{T}$ by which sellers update their beliefs as in (7), equilibrium consists of

(i) the threshold bidding rule as defined in 3.1.i

(ii) the bidder arrival parameter δ_0 that determines mean bidder arrival $\Lambda(r \mid \delta_0)$ as in (3)

(iii) the seller reserve price rule defined by $\psi(r^* \mid b, c) = v_0$ for any seller beliefs b

(iv) the seller entry threshold $\bar{v}(b, c)$ implied by (8)

such that actual bidders maximize expected profit from participating in an auction, potential bidders earn zero expected profit from entering any auction, sellers maximize profits conditional on their beliefs about the bidder entry process, and the marginal seller earns zero expected profit given their beliefs about the bidder entry process.

The equilibrium conditions highlight the importance of seller beliefs in determining outcomes on the auction platform. Sellers' beliefs determine both their entry decision and their choice of reserve price conditional on entry. Further, present beliefs affect future beliefs through the updating rule $\mathcal{T}$, so the entire path of sellers' entry decision and reserve prices depend on their initial beliefs. While bidders' equilibrium strategies do not depend on sellers' beliefs about the bidder arrival process, bidder choices and outcomes are best responses to sellers' behavior. Thus, all actions on the platform are shaped by sellers' beliefs about bidder behavior, and the speed with which they learn about the true parameters.

4 Estimation

The model from the previous section highlights several important features of the sellers' environment. First, sellers' beliefs determine both their decision of whether to list an item and, conditional on listing, what reserve price to set. Further, the platform setting implies that profit-maximizing sellers with full information will set a lower reserve price than would be optimal if entry did not depend on sellers' reserve prices. I now estimate the model to evaluate the role these features play in the empirical patterns, and the extent to which sellers are learning the true bidder arrival process.

Estimation proceeds in two parts. First, I estimate the demand side of the model, consisting of bidder arrival parameters, the distribution of bidder valuations F_B, and heterogeneity in mean item valuations. To do this, I derive a likelihood function for the demand side that corrects for selection in which bids are observed due to the increasing minimum bid rule, and apply results on debiased machine learning to flexibly estimate observed item heterogeneity. This process also implies bidder time costs from bidders' zero profit condition. I then estimate the supply side of the model, using data from both experienced and inexperienced sellers. I recover the distribution of seller values by assuming experienced sellers have accurate beliefs about the bidder arrival process, and obtain time cost estimates from the seller entry condition. I then use the implied seller value distribution to estimate new sellers' prior beliefs about the bidder arrival process and study the relationship between seller learning and selection.

4.1 Demand side

I first estimate the parameters of the demand model, which include both a low-dimensional component and a high-dimensional component. The low-dimensional parameter ϑ_d consists of the bidder arrival parameters, and the bidder value distribution F_B and reserve price distribution F_r (each approximated using Gaussian mixture models with five components). The high-dimensional parameter γ accounts for observable heterogeneity in item values through an index $\gamma(X_j)$ for each item j, where X_j is text data that is encoded as a vector of 5,368 word indicator variables from the item description as well as month fixed effects. In order to flexibly account for this heterogeneity, I parameterize γ as a large neural network. Since machine learning models can introduce bias due to both overfitting and model selection of high-dimensional parameters, I use debiased machine learning techniques as in Farrell, Liang, and Misra (2020), Chernozhukov et al. (2022), and Ichimura and Newey (2022) to correct for bias in ϑ_d. More details on the data construction and estimation process may be found in Appendix A3, and details on the orthogonalization process are shown in Appendix A4.

In addition to accounting for item-level heterogeneity in mean valuations, I also allow bidders' entry equilibrium to vary with observable seller and item characteristics. I parameterize the number of bidders to enter an auction as

$$\Lambda_j \equiv \Lambda(r_j, Z_j \mid \lambda, \delta_0) = \exp(\delta_{0,1} + Z_j'\lambda + \delta_{0,2}\rho(r_j)) \tag{9}$$

where Z_j contains the seller feedback score, seller ratings, and the average log item value $\gamma(X_j)$. I assume the additional arrival parameter λ is known to all sellers, regardless of their level of experience. To streamline notation, I combine the known arrival shifter $Z_j'\lambda$ with the intercept to write the item-specific arrival parameter $\delta_{j,0,1} \equiv \delta_{0,1} + Z_j'\lambda$, with $\delta_{j,0} \equiv \{\delta_{j,0,1},\ \delta_{0,2}\}$. I also use this notation for beliefs, where b_j represents sellers' item-specific beliefs about bidder arrival with $\mathbb{E}_{b_j}[\delta_1] = \mathbb{E}_b[\delta_1] + Z_j'\lambda$.

The function ρ is determined by the bidder zero-profit condition at each point in the support of r. Since evaluating the high-dimensional index γ is computationally challenging, I opt for a reduced-form representation of the entry process instead of solving the fixed-point problem for ρ. I set ρ to be the product of the identity function and the CDF of the bidder value distribution evaluated at r, i.e. $\rho(r) = rF_B(r)$. The intuition for this choice comes from the zero profit condition for bidder entry and Proposition 1. The expected bidder surplus from entering the auction depends on the probability that their bid will be below the reserve price, $F_B(r)$; it is also strictly decreasing in r, which is the price paid when the reserve is binding. I also estimate the model with alternative specifications for ρ and find similar results.

eBay's use of an increasing minimum bid creates a selection problem, since only the first two highest bids in an auction are known to correspond to the first two highest-value bidders. All other bids may be excluded from the data if the two highest bids are the first to be placed, since the minimum bid will increase to the second highest of these two and "lock out" subsequent arrivals (Platt 2017; Freyberger and Larsen 2022). To counteract this problem, I derive a likelihood that explicitly models the selection process (see Appendix A5 for a detailed derivation). I first denote $v_j^{(k)}$ as the kth highest homogenized log-bid. Using the Poisson assumption on bidder arrival, I model the distributions of the reserve price r_j and the highest two bids $v_j^{(k)}$ when these bids are observed.[6] This likelihood also conditions on the starting minimum bid m_j (which is binding when there are fewer than two bids), the number N_j of bids observed, arrival shifters Z_j, and components of observable heterogeneity X_j. Together, the demand likelihood contribution for a single auction

[6] Note that r_j and $v_j^{(k)}$ are not observed directly; rather, each is a residual representing its counterpart in the data after being homogenized using the item value index $\gamma(X_j)$.

is

$$
\begin{aligned}
\ell_j^{\text{demand}}(\vartheta_d, \gamma \mid N_j, \{v_j^{(k)}\}, X_j, m_j, r_j) = \\
& f_r(r_j \mid \vartheta_d, \gamma) \cdot \left[e^{-\Lambda[1-F_B(m_j|\vartheta_d,\gamma)]}\right]^{\mathbb{1}[N_j=0]} \\
& \cdot \left[f_B(v_j^{(1)} \mid \vartheta_d, \gamma)\Lambda e^{-\Lambda_j[1-F_B(m_j|\vartheta_d,\gamma)]}\right]^{\mathbb{1}[N_j=1]} \\
& \cdot \left[f_B(v_j^{(1)} \mid \vartheta_d, \gamma) f_B(v_j^{(2)} \mid \vartheta_d, \gamma)\Lambda_j^2 e^{-\Lambda_j[1-F_B(v_j^{(2)}|\vartheta_d,\gamma)]}\right]^{\mathbb{1}[N_j\geq 2]}
\end{aligned}
\tag{10}
$$

Appendix A5 shows how this likelihood approach performs with simulated data. In order to illustrate the role of the orthogonalization step in consistently estimating ϑ_d, I simulate settings with a low-dimensional γ and a high-dimensional γ.

After estimating the bidder value distribution F_B and the other bidder arrival parameters, I estimate the bidder entry cost c_B^E. For each auction, I compute the ex ante expected surplus $\pi_B(r \mid \Lambda, c)$ and use the bidder zero-profit condition in (2) to estimate bidders' entry cost as the mean of the expected bidder surplus across all auctions. Since bidder insertion fees c_B^I are zero in the dataset, this implies that the full bidder entry cost is in fact the time cost c_B^T.

4.2 Supply side

I now turn to the problem of identifying and estimating the supply side of the model. Since all demand-side parameters are recovered from the likelihood approach in the previous section, the remaining parameters of interest are the distribution F_S of sellers' outside option values, sellers' entry cost distribution $F_{c_S^E}$, and new sellers' beliefs about the unknown bidder arrival parameter δ_0. While seller values and costs have been estimated in many similar settings, the belief estimation process is key to understanding the role of information and learning in determining outcomes on the auction platform.

I first estimate all seller parameters except new seller beliefs using data from experienced sellers. I denote these parameters, including the nuisance parameters of distribution of entry parameters $F_{\delta_{0,1}}$, by ϑ_s. I make the simplifying assumption that the most experienced sellers have perfect information about the arrival process, so plugging the experienced sellers' reserve prices into the virtual type function $\psi(\cdot \mid \delta_0, c)$ yields the imputed seller values $\hat{v}_0$ for all items that are listed for auction. I also assume sellers' entry costs are i.i.d. Exponential, their values v_0 are i.i.d. from a

5-component Gaussian mixture, and entry parameters $\delta_{j,0,1}$ are i.i.d. Gaussian. First denote $\tilde{c}(z)$ as the vector of costs where c_S^E is replaced with z and $\Pi^*(v_0 \mid \delta_0, c)$ as the maximized profit given seller value v_0. Further, denote by $\bar{v}(\delta_{j,0}, c)$ the experienced seller entry threshold with known bidder entry parameter $\delta_{j,0}$ and costs c. The likelihood contribution of a single auction run by an experienced seller is

$$\ell_j^{\text{supply-e}}(\vartheta_s \mid \hat{v}_{0j}, \delta_{j,0}) = \frac{f_s(\hat{v}_{0j} \mid \vartheta_s) \cdot f_{\delta_{0,1}}(\delta_{j,0} \mid \vartheta_s) \cdot F_{c_S^E}(\Pi^*(\hat{v}_{0j} \mid \delta_{j,0}, \tilde{c}(0)) - \hat{v}_{0j} \mid \vartheta_s)}{\int_0^\infty \left[\int_{-\infty}^\infty F_S(\bar{v}(l, \tilde{c}(z)) \mid \vartheta_s) f_{\delta_{0,1}}(l \mid \vartheta_s) dl \right] f_{c_S^E}(z \mid \vartheta_s) dz}$$

The density for seller values is multiplied by a correction term to account for random truncation in the seller value distribution from the seller threshold rule. That is, each item is only listed if the entry cost is below the expected surplus of the seller with a particularly favorable (zero) entry cost. The probability that this occurs for a specific imputed item value $\hat{v}_{0j}$ is divided by the probability that any seller lists their item. Due to seller entry costs being unobserved, this likelihood is not identified on its own. To identify the distribution of seller values, I assume that the mean of the seller entry cost distribution is equal to c_S^P, where the platform entry fees are observed in the data. Variation in $\delta_{j,0,1}$ shifts sellers' participation threshold and traces out the distribution of seller values.

Having obtained the sellers' value distribution, I can use variation in new sellers' reserve prices to identify their underlying beliefs in each period. The following proposition offers conditions under which seller beliefs b_{δ_2} about the effect of reserve prices on arrival can be semiparametrically identified using variation in the highest bid $v^{(1)}$.

Proposition 5. Denote history $\mathcal{H}$ as a collection of data $\mathbf{D}$ from auctions. Let Assumption 1 hold, assume F_S is known, and further assume

(i) The prior b is common to all sellers with a shared history $\mathcal{H}$.

(ii) b is composed of independent marginal densities b_{δ_1} and b_{δ_2}, where both the marginal b_{δ_1} corresponding to the intercept and the seller entry threshold $\bar{v}(b, c)$ are known.

(iii) b_{δ_2} satisfies the Carleman condition, i.e. the absolute moments of b_{δ_2} (written as $\mu_j = \int |\delta|^j b_{\delta_2}(\delta_2) d\delta_2$) are finite for all $j \geq 1$ and satisfy $\sum_{j=1}^\infty \mu_j^{-1/j} = \infty$.

(iv) The reserve price r has full support on a positive-measure interval including some x^* for which $\rho(x^*) = 0$ and $\psi(x^* \mid b, c) < \bar{v}(b, c)$.

(v) For $\xi_k(\beta, x^*, c)$ defined as $\frac{\partial^{k-1}}{\partial x^{k-1}}\psi(x^* \mid b, c)$ with all terms of the form $\sum_{n=0}^{\infty} G_n(x^*)\frac{\partial^k}{\partial x^k}p_n(x^* \mid b)$ replaced with $\beta \sum_{n=0}^{\infty} G_n(x^*)\frac{\partial^k}{\partial x^k}p_n(x^* \mid b)$ for $G \in \{R, K\}$, ξ_k is invertible in β.

Then the marginal density b_{δ_2} of a seller with history $\mathcal{H}$ is identified up to its first $\bar{k}$ moments. (Proof in Appendix A2)

The proof proceeds in two parts, for which I give a heuristic explanation here. The first step is to recover the virtual type function ψ that sellers use to determine their choice of reserve price. In the absence of seller selection, the distribution of sellers' reserve prices is obtained via a standard change-of-variables approach using the first-order condition $v_0 = \psi(r \mid b, c)$. Combined with knowledge of sellers' participation threshold $\bar{v}(b, c)$, the distribution of reserve prices for the listed items can be written as a known, invertible function of $\psi(\cdot \mid b, c)$.

The next step of the proof is to invert the virtual type function to recover the marginal density b_{δ_2}, which represents seller beliefs about the effect of the reserve price on bidder arrival. Sellers' first-order conditions are composed of integrals of known functions, making this inversion similar to the identification of distributions of random coefficients (Fox et al. 2012). In this case, $\psi(\cdot \mid b, c)$ and its derivatives are known functions of various moments of b. Additional restrictions on b (namely, independence between the marginal beliefs about $\delta_{0,1}$ and $\delta_{0,2}$ and knowledge of sellers' beliefs about $\delta_{0,1}$) simplify the integrals in question. Thus, evaluating the virtual type function and its derivatives yields known, invertible functions of moments of b_{δ_2}. The identification result is general in that it recovers an arbitrary number of moments of the marginal belief distribution (allowing for a broad class of possible beliefs), though it is important to note that these beliefs are only identified within the context of the larger parametric model. Notably, it is not necessary to assume that sellers are Bayesian: identification of b_{δ_2} is achieved with data from a single period for all sellers that observe the same data history $\mathcal{H}$.

Though Proposition 5 offers semiparametric identification of b_{δ_2}, the necessary assumptions are somewhat restrictive. First, Assumption 1 imposes that all sellers' outside option values are drawn i.i.d. from the same distribution, which rules out time-invariant heterogeneity in sellers' value distributions. Also, assumption (i) of the proposition rules out unobserved determinants of beliefs: given the form of the updating rule in equation (7), this allows the researcher to restrict attention to sellers with identical beliefs. Further, since only one signal (profit) is observed after every auction, I impose independence between the two marginal densities of the prior distribution. The second assumption is also restrictive: other components of the sellers' decision problem must be known to isolate the effect of beliefs about any one parameter. This is similar to functional form assumptions of e.g. an additive T1EV shock in logit demand models, even while the rest of a utility specification may be flexible. Assumptions (iii) and (iv) are similar to assumptions in the random coefficients literature, principally Fox et al. 2012; these use variation in a linear index to recover population densities, though in this case I study an individual's belief density. Assumption (v) is also technical, and requires the derivatives of the virtual type function to be invertible in weighted sums of the derivatives of the expected probability of any bidder arriving. Since the derivatives of the virtual type function are known, this is a joint restriction on the seller beliefs about δ_1 and the bidder value distribution F_B at the point x^* in assumption (iv).

This result is related to others in the literature on identifying individual beliefs in structural models. Lu (2019) shows that state-dependent beliefs can be identified in a setting with finite support and Bayesian updating; in contrast, I do not require Bayesian updating and allow for absolutely continuous density functions. Wang et al. (2024) likewise adopts a finite-support approach with Bayesian updating, which is used to identify beliefs about time-varying, unobserved macroeconomic trends; Wang and Yang (2024) offers more general results in finite-support settings for both myopic and forward-looking agents. Aguirregabiria and Magesan (2020) semiparametrically identifies firm beliefs within a game, and similarly relies on a finite support. While they do not require Bayesian updating, they require beliefs to coincide with the truth in some cases; this restriction on seller beliefs functions similarly to my assumption that other features of sellers' decision problems are pinned down by external arguments.

Though beliefs are semiparametrically identified for each history $\mathcal{H}$ of auction data, in practice I make several additional assumptions for computational tractability. First, I assume sellers' initial beliefs about δ_0 are bivariate normal, with parameters jointly denoted as ϑ_b. I also assume beliefs are updated according to a modified Laplace approximation to Bayes' rule: each period, sellers' beliefs are a bivariate normal with mean equal to the maximum a posteriori estimate of the true Bayesian posterior and covariance matrix given by the curvature of the true Bayesian posterior at the maximum a posteriori estimate.[7] I use $\psi(r^* \mid b, c)$ as a control function for seller values v_0 since they are an unobservable but critical component of sellers' updating process. Finally, I use these parametric assumptions on the entry cost distribution and sellers' prior to pin down sellers' entry threshold $\bar{v}(b, c)$ and prior b_{δ_1}, which allows me to relax the independence assumption in Proposition 5. Appendix A6 discusses additional details of the estimation procedure.

Having established the identification of new sellers' prior parameters, I now explain the likelihood approach used for estimation. I use a change of variables to obtain the density of the new sellers' reserve prices from their underlying value distribution. As with the experienced-seller likelihood, there is a selection term to account for the probability of each item being listed by any given seller. The likelihood contribution of a single auction of item j run by an inexperienced seller in their tth auction is therefore

$$\ell_{jt}^{\text{supply-i}}(\vartheta_b \mid \mathbf{D}_{jt}) = f_s(\hat{v}_{0j} \mid \vartheta_s) \cdot \psi'(r_j \mid b_t, \tilde{c}(0)) \cdot \frac{F_{c_S^E}(\Pi^*(\hat{v}_{0j} \mid b_t, \tilde{c}(0)) - \hat{v}_{0j} \mid \vartheta_s)}{\int_0^\infty F_S(\bar{v}(b_t, \tilde{c}(z)) \mid \vartheta_s) \cdot f_{c_S^E}(z \mid \vartheta_s) dz}$$

$$\begin{aligned} s.t. \quad & b_{t+1} = \mathcal{T}(b_t, \mathbf{D}_{jt} \mid \vartheta_b) \ \forall t \\ & \hat{v}_{0j} = \psi(r_j \mid b_t, \tilde{c}(0)) \end{aligned}$$

Though beliefs in this model can be identified from data within a single period, additional variation across and within different sellers is pooled via the assumed learning rule to aid in estimation.

4.3 Estimates

Table 1 presents the estimated bidder arrival parameters. As predicted by the model, the reserve price coefficient $\delta_{0,2}$ is negative, indicating that higher reserve prices deter potential bidders from entering auctions. Arrival also increases with reputation variables like seller rating and feedback, as well as the estimated mean item value. The signs of the estimates are intuitive: more bidders enter auctions when the sellers have higher ratings, are more experienced, and when items are more

[7] This assumption helps solve a computational challenge, since the prior is not conjugate with the posterior due to the nonlinear dependence of expected profit on the bidder arrival parameters. In addition to this nonlinearity, each sellers' path of beliefs about the arrival parameter can evolve differently according to their signals. The Laplace approximation is one of several posterior approximations used in Bayesian statistics, and imposing that beliefs are updated in this manner ensures that learning follows a computationally tractable Markov process with a relatively low-dimensional state.

valuable.

Table 1: Estimated bidder arrival parameters and entry cost

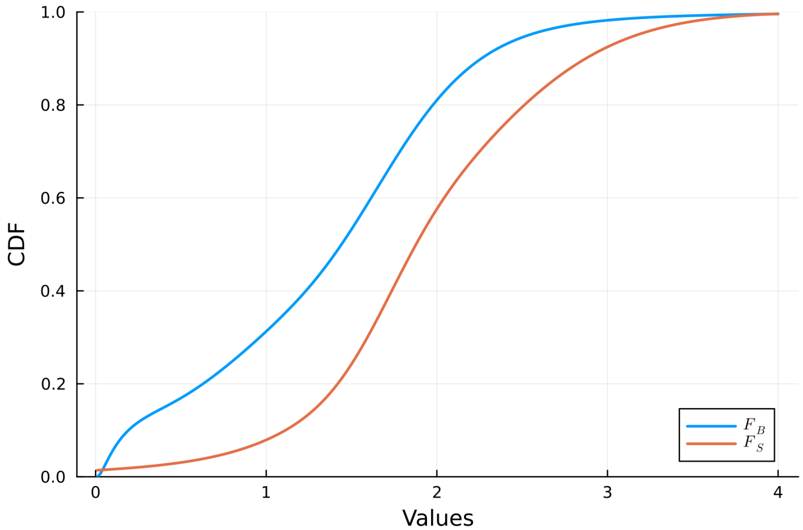
			λ_0			
	$\delta_{0,1}$	$\delta_{0,2}$	Rating	IHS(Feedback)	ln(Pred. Item Value)	c_B^T
Estimate	0.871	-0.245	0.055	0.146	0.409	0.056
Std. Err.	(0.059)	(0.007)	(0.026)	(0.010)	(0.022)	(5.5×10^{-5})

Notes: The estimated coefficients are obtained via debiased GMM; further details are presented in Appendix A4. Estimates and standard errors for all but the entry cost are computed using the optimal weighting matrix; the standard error for the entry cost is the naive standard error (*standard error correction for two-step estimation in progress*).

The estimated bidder entry cost comes from the equilibrium bidder entry condition and the other estimated bidder arrival parameters. Evaluating the expected zero profit condition in equation (2) yields an estimated homogenized bidder time cost c_B^T of 0.056, since there are no bidder insertion fees. For the average (median) item in the dataset, this is approximately \$0.48 (\$0.33). This represents a moderate but not prohibitive time cost to entering each auction and inspecting the listing. The average homogenized seller entry fee is $c_S^P = 0.075$, though sellers' overall entry costs are assumed to be heterogeneous for different items.

As described in the previous section and as is common in the auction literature, the estimated demand parameters yield the optimal reserve pricing rule for sellers with perfect information. I use the sample of reserve prices chosen by experienced sellers (defined as those in the top 25% of sellers by experience at the start of the data) to impute the sellers' outside option for each item, under the assumption that experienced sellers have perfect information about the bidder arrival process. Figure 6 plots the estimated value distribution F_S along with the bidder value distribution F_B obtained in the demand-side estimation.

Figure 6: Estimated value distributions for auction participants



Notes: The bidder value distribution F_B is estimated via the maximum likelihood approach in 4.1 using a 5-component Gaussian mixture model for log values. The seller outside option distribution F_S is fit to imputed seller values among experienced sellers, as in 4.2, and uses a 5-component Gaussian mixture model for seller values.

The estimated seller value distribution largely first-order stochastically dominates the estimated bidder value distribution. Since the population of sellers is the group of users who have previously acquired Beanie Babies, it is reasonable for them to have a higher value distribution for these items than any random bidder. However, this difference in value distributions is not unreasonably large: the seller value distribution largely falls between distributions of the maximum value of two bidders and that of three bidders (which are not plotted here), so it may be profitable in expectation for a seller with a large $\hat{v}_{0j}$ to list an item for sale.

I now turn to the estimates of new sellers' beliefs and their learning process. I estimate the model on a sample of all sellers with at least 5 auctions, to have a population of "serious" sellers who have more than a couple items to sell. I also limit the sample to the first 5 auctions of all such sellers to not bias the estimates with subsequent exit of some of these sellers. The estimated prior parameters are shown in Table 2, for two cases where only the arrival coefficient $\delta_{0,2}$ is unknown to new sellers (version (a)) and for the general case where both parameters in δ_0 are unknown (version (b)). In both cases, the prior mean for the reserve price coefficient $\delta_{0,2}$ is higher than the estimated parameter -0.245, implying that new sellers' beliefs about bidder arrival are upwardly biased, and particularly so for auctions with high reserve prices. The marginal beliefs about $\delta_{0,2}$ are similar across both specifications, with a moderately dispersed prior that allows for some learning, though it is far from immediate.

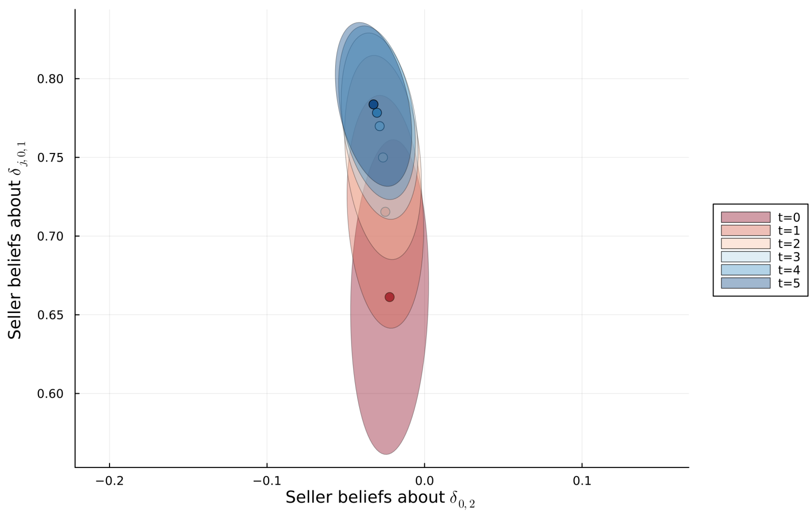
Table 2: Estimated new seller priors about the bidder arrival process

	(a)	(b)	
Parameter	$\delta_{0,2}$	$\delta_{0,1}$	$\delta_{0,2}$
Prior Mean	-0.042	0.661	-0.022
	(1.7e-5)	(3.2e-5)	(1.4e-5)
Prior Std. Dev.	0.317	1.0	0.248
	(2.6e-5)	(4.6e-5)	(1.4e-5)
Prior Correlation	-	0.091	-
		(1.9e-5)	

Notes: The model is estimated on the first 5 auctions of all 3,975 new sellers that list at least 5 auctions for sale. Version (a) treats the intercept parameter $\delta_{0,1}$ as known by all new sellers, so the only uncertainty is about the arrival coefficient. Version (b) treats both parameters as unknown to new sellers. For this approximation, prior standard deviations are bounded from above by 1. Standard errors are naive standard errors, treating seller value distribution parameters as known (*standard error correction for two-step estimation in progress*).

To help interpret the estimated prior parameters in Table 2, I simulate the path of new sellers' average beliefs about the unknown parameter δ_0 for the case where both parameters are unknown. Figure 7 plots ellipses corresponding to the estimated beliefs of new sellers, though the contours correspond to a 0.1-standard deviation interval (rather than a conventionally-sized credible interval) for clarity. The prior mean drifts toward the true parameters as new sellers learn from successive auctions. While most of the shift is due to an upward correction in seller beliefs about $\delta_{0,1}$, the posterior mean begins to curve toward the true parameter δ_0.

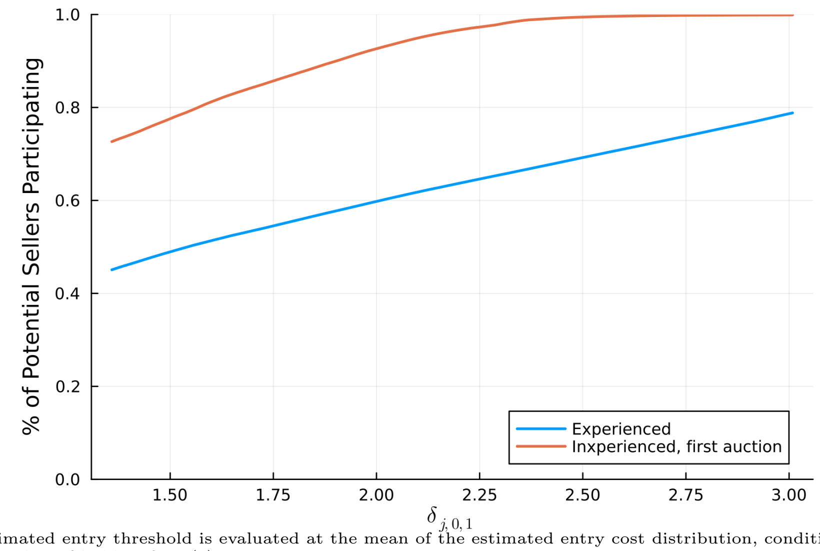
Figure 7: Estimated path of average new seller beliefs



Notes: The ellipses represent the average beliefs of new sellers in their first 5 auctions, conditioning on all 3,975 new sellers who list at least 5 auctions. The contours represent a 0.1-standard deviation around the mean of the belief distribution.

Finally, I directly show the effect of seller beliefs on selective entry by plotting the average seller entry threshold for new sellers relative to the experienced seller entry threshold. Figure 8 shows how the experienced sellers' entry threshold is increasing in the expected arrival rate, where the entry threshold is evaluated at the mean of the estimated entry cost distribution $F_{c_S^E}$. The entry threshold is significantly higher among new sellers in their first auction. This is consistent with Table 2: in spite of the lower prior mean for $\delta_{0,1}$ (relative to the true parameter), its high variance combined with the higher prior mean for $\delta_{0,2}$ makes entry attractive to new sellers. This relative ordering of entry thresholds is consistent with the selection pattern in Figure 4, where new sellers that exit early also set higher reserve prices in their first auctions.

Figure 8: Estimated average seller entry threshold conditional on baseline bidder arrival parameter $\delta_{j,0,1}$



Notes: The estimated entry threshold is evaluated at the mean of the estimated entry cost distribution, conditional on the log of the expected number of bidders for $\rho(r) = 0$.

5 Platform information problem

I now use the estimated model to study the problem of information provision by the platform. There is a unit mass of sellers on the platform, where ω is the share of new sellers with beliefs b_0 about the bidder arrival process, and $1 - \omega$ is the share of experienced sellers with accurate beliefs about the bidder arrival process. For simplicity, I treat this as a one-stage problem where each seller lists a single item and no additional sellers join the platform.

I assume the platform can choose the seller-facing fees c_S^P and c_S^E as well as the number of auctions a in a dataset $\mathbf{D}_a$ that it shows to new sellers before they list their first item.[8] The platform can only truthfully reveal information from auctions that have previously occurred, and therefore provides a random subsample $\mathbf{D}_a$ of recent auction data to new sellers to update their priors from b_0 to $\mathcal{T}(b_0, \mathbf{D}_a)$. I use the shorthand $a = \infty$ for the platform sharing all of its data with new sellers; this is equivalent to the platform providing an automated pricing tool to all sellers. This data is drawn i.i.d. from the distribution $F_{\mathbf{D}}$ of all auction data owned by the platform. Formally,

[8]I treat bidder-facing fees c_B^P as being fixed at zero. This is motivated both by the fee being equal to zero for bidders in the data, as well as analysis by Marra (2019) on a wine auction platform indicating that increasing revenue while maintaining transaction volume requires setting $c_B^P < 0$.

the platform's profit maximization problem is

$$\max_{c_S^P,\, c_S^E,\, a} \int\int\Bigg(\omega\int \underbrace{F_S(\bar{v}(\mathcal{T}(b_0,\mathbf{D}_a),\tilde{c}(z)))}_{\mathbb{P}[\text{Entry} \mid \text{Inexperienced}]}\cdot \underbrace{\left[c_S^E + c_S^P\cdot\int R(r^*(v_0\mid\delta,c_S^P)\mid\delta_0)\mathcal{T}(b_0(\delta),\mathbf{D}_a)d\delta\right]}_{\text{Platform revenue} \mid \text{entry}} dF_{\mathbf{D}}(\mathbf{D}_a)$$
$$+ (1-\omega)\cdot \underbrace{F_S(\bar{v}(\delta_0,\tilde{c}(z)))}_{\mathbb{P}[\text{Entry} \mid \text{Experienced}]}\cdot\underbrace{\left[c_S^E + c_S^P\cdot R(r^*(v_0\mid\delta_0,c_S^P)\mid\delta_0)\right]}_{\text{Platform revenue} \mid \text{entry}}\Bigg)dF_S(v_0)dF_{c_S^E}(z) - c_a^P\cdot a \tag{11}$$

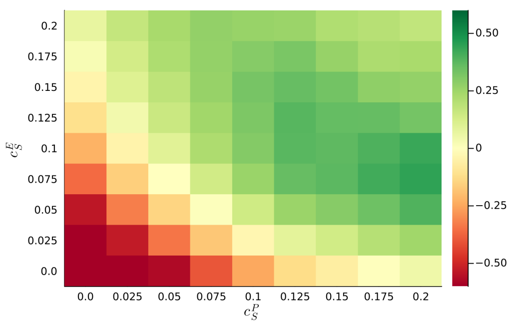
where c_a^P is the marginal cost of providing data from a auctions to each seller, and $r^*(v_0 \mid \delta, c_S^P)$ is the optimal reserve price function conditional on sellers' outside option v_0, the arrival parameter δ, and the sellers' revenue fee. Throughout this exercise I set $c_a^P = 0$ to reflect the low marginal cost of providing additional data, assuming it has already been collected. Though it is likely that there is significant fixed cost in providing sellers with data, I assume this cost is sunk and not relevant for the platform's problem.

The maximization problem in (11) highlights the importance of sellers' beliefs in the platform's optimal information problem. The first line corresponds to the new sellers on the platform, each of whom observes a dataset $\mathbf{D}_a$ with a auctions before listing their first item. Each seller only chooses to list an item if their updated beliefs $\mathcal{T}(b_0, \mathbf{D}_a)$ about the bidder entry process imply it is optimal to do so; conditional on entry, they will also choose a reserve price according to their beliefs. Thus, knowledge of b_0 is critical for understanding how any additional information $\mathbf{D}_a$ will shift seller beliefs and therefore behavior. The second line corresponds to experienced sellers, whose selective entry and choice of reserve price are informed by their perfect information about the bidder entry process. While these sellers are not directly affected by the platform's information provision, they must pay the revenue fee c_S^P which is chosen by the platform jointly with its dataset length a. Thus, the fee paid by experienced sellers is influenced by new sellers' beliefs about bidder entry and how much information new sellers receive from the platform.

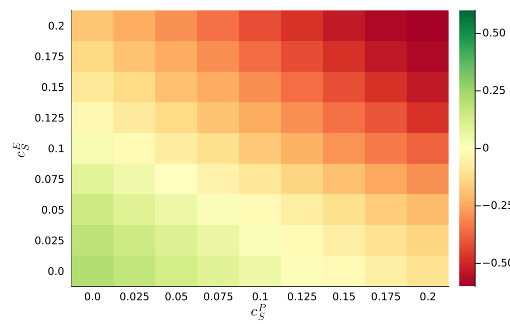
I now compute the expected platform profit for different c_S^P, c_S^E, and a. To fix ideas, I set $\omega = 0.5$ as the fraction of sellers who are new, so a large portion of users may be affected by the platform providing additional information. In Figure 9, I plot the estimated effect of alternative fee structures without information provision (i.e., $a = 0$) on various outcomes of interest. As may be expected, the largely platform gains from higher fees while the sellers would prefer lower fees.

Also, higher fees reduce the probability that sellers list items on the platform, which contributes to the lower number of bidders.

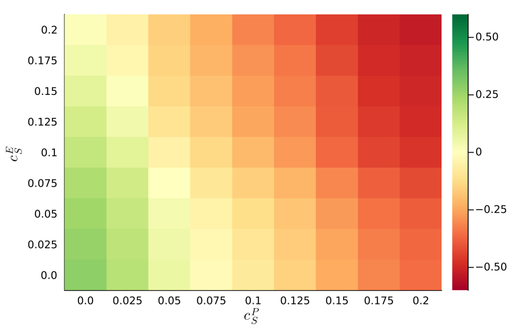
Figure 9: Percent changes in outcomes under alternative fee structures, $a = 0$



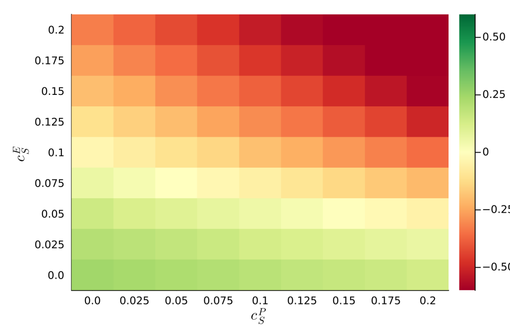
(a) Platform revenue



(b) Number of bidders



(c) Seller surplus



(d) Seller participation rate

Notes: Each scenario was evaluated form a grid of possible fees, simulating 1,000 sellers for each combination of parameter values. Green (red) squares represent a percentage increase (decrease) in a given outcome relative to the outcome under the baseline values of c_S^P and c_S^E; each panel depicts changes between -60% and 60%.

To highlight the role of information provision in this setting, I contrast the benchmark of $a = 0$ with $a = \infty$: the platform provides all information about the data-generating process to sellers. Table 3 compares outcomes under the two information regimes. The first row, which is Pareto optimal for $a = 0$, is the baseline fee structure in the data.[9] The Pareto optimal fee structure under $a = \infty$ yields strictly better platform and seller outcomes, and even gives a marginal increase in the number of bidders using the platform.

[9]Since bidders have a zero profit condition that determines entry, in expectation no bidder's welfare will be changed by these regimes. To compare bidder outcomes to those of sellers and the platform, I measure the average number of bidders that enter an auction on the platform. Any "Pareto" improvement in this exercise must maintain at least the same number of bidders as under the baseline fee structure.

Table 3: Platform-optimal and Pareto-optimal fees under different information provision

a	Objective	c_S^P	c_S^E	Platform revenue	Seller surplus	Number of bidders
0	Pareto	0.05	0.075	0.093	0.237	4.01
0	Platform	0.2	0.075	0.134	0.135	2.82
∞	Pareto	0.125	0.025	0.102	0.243	4.05
∞	Platform	0.2	0.05	0.128	0.202	3.14

Notes: Platform revenues and seller surplus are measured per potential seller. The two objectives are "Platform" (maximizing platform revenues) and "Pareto", which I define to be the platform-revenue maximizing fee structure such that platform revenues, seller surplus, and the number of bidders are at least as great as the benchmark case.

Interestingly, platform revenue is highest in the case where $a = 0$, or the platform shares no information with new sellers. Since uninformed sellers are on average more willing to list an item for auction (see Figure 8), the platform can increase entry fees relative to the full-information setting. While this exercise abstracts from dynamic considerations, including both subsequent learning by existing sellers and additional entry by new users, it suggests that full, free information will not always be chosen by profit-maximizing platforms. Despite platforms' incentives to serve as a matching mechanism and facilitate a large volume of transactions, some user uncertainty may be optimal.

6 Conclusions

This paper studies the problem of information provision by an auction platform where sellers face uncertainty about the bidder arrival process. I first present evidence that new sellers learn to set optimal reserve prices as they gain more experience. I pair a model of seller learning with a model of two-sided endogenous entry onto an auction platform to investigate how new seller behavior is driven by both selection and learning. The model implies a new reserve price formula that is designed to both attract bidders to the auction and extract surplus from them; failure to account for the negative effect of high reserve prices on bidder entry leads new sellers to set higher-than-optimal reserve prices. I show that sellers' beliefs about the bidder arrival process can be semiparametrically identified from reserve price data under certain conditions, and estimate new sellers' beliefs.

These results highlight platforms' ability to influence user behavior outside of its well-known ability to charge different fees to different sides of the market. Information provision improves the quality of the marketplace for sellers, who are able to better optimize their entry decisions and pricing strategy. Sellers' improved profits also increase platform profits through increased revenues, and induce additional bidder entry through lower average prices on the platform. However, providing or witholding information may enhance the platform's ability to extract revenue through alternative fee structures.

References

Aguirregabiria, Victor, and Arvind Magesan. 2020. "Identification and estimation of dynamic games when players' beliefs are not in equilibrium". *The Review of Economic Studies* 87 (2): 582–625.

Bronnenberg, Bart J, Jun B Kim, and Carl F Mela. 2016. "Zooming in on choice: How do consumers search for cameras online?" *Marketing science* 35 (5): 693–712.

Bulow, Jeremy, and Paul Klemperer. 1996. "Auctions versus Negotiations". *American Economic Review* 86 (1): 180–94.

Chernozhukov, Victor, et al. 2022. "Locally robust semiparametric estimation". *Econometrica* 90 (4): 1501–1535.

Cogley, Timothy, Riccardo Colacito, and Thomas J Sargent. 2007. "Benefits from US monetary policy experimentation in the days of Samuelson and Solow and Lucas". *Journal of Money, Credit and Banking* 39:67–99.

Compiani, Giovanni, et al. 2022. *Online Search and Product Rankings: A Double Logit Approach.* Tech. rep. Working paper.

Davis, Andrew M, Elena Katok, and Anthony M Kwasnica. 2011. "Do auctioneers pick optimal reserve prices?" *Management Science* 57 (1): 177–192.

Engelbrecht-Wiggans, Richard. 1987. "On optimal reservation prices in auctions". *Management Science* 33 (6): 763–770.

Erdem, Tülin, and Michael P Keane. 1996. "Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets". *Marketing science* 15 (1): 1–20.

Farrell, Max H, Tengyuan Liang, and Sanjog Misra. 2020. "Deep learning for individual heterogeneity: An automatic inference framework". *arXiv preprint arXiv:2010.14694*.

Foroughifar, Mohsen. 2023. "The challenges of deploying an algorithmic pricing tool: Evidence from airbnb". PhD thesis, University of Toronto (Canada).

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2016. "The slow growth of new plants: Learning about demand?" *Economica* 83 (329): 91–129.

Fox, Jeremy T, et al. 2012. “The random coefficients logit model is identified”. *Journal of Econometrics* 166 (2): 204–212.

Freyberger, Joachim, and Bradley J Larsen. 2022. “Identification in ascending auctions, with an application to digital rights management”. *Quantitative Economics* 13 (2): 505–543.

Gomes, Renato. 2014. “Optimal auction design in two-sided markets”. *The RAND Journal of Economics* 45 (2): 248–272.

Haggag, Kareem, Brian McManus, and Giovanni Paci. 2017. “Learning by driving: Productivity improvements by new york city taxi drivers”. *American Economic Journal: Applied Economics* 9 (1): 70–95.

Hasker, Kevin, and Robin Sickles. 2010. “eBay in the economic literature: Analysis of an auction marketplace”. *Review of Industrial Organization* 37 (1): 3–42.

Hitsch, Günter J. 2006. “An empirical model of optimal dynamic product launch and exit under demand uncertainty”. *Marketing Science* 25 (1): 25–50.

Hodgson, Charles, and Gregory Lewis. 2023. *You can lead a horse to water: Spatial learning and path dependence in consumer search.* Tech. rep. National Bureau of Economic Research.

Huang, Yufeng, Paul B Ellickson, and Mitchell J Lovett. 2020. “Learning to set prices”. *Available at SSRN 3267701.*

Ichimura, Hidehiko, and Whitney K Newey. 2022. “The influence function of semiparametric estimators”. *Quantitative Economics* 13 (1): 29–61.

Jullien, Bruno, Alessandro Pavan, and Marc Rysman. 2021. “Two-sided markets, pricing, and network effects”. In *Handbook of Industrial Organization*, 4:485–592. 1. Elsevier.

Katkar, Rama, and David H Reiley. 2007. “Public versus secret reserve prices in eBay auctions: Results from a pokémon field experiment”. *The BE Journal of Economic Analysis & Policy* 6 (2): 0000102202153806371442.

Keller, Godfrey, and Sven Rady. 1999. “Optimal experimentation in a changing environment”. *The review of economic studies* 66 (3): 475–507.

Kim, Yewon. 2020. “Customer retention under imperfect information”. PhD thesis, The University of Chicago.

Klein, Benjamin, et al. 2005. "Competition in two-sided markets: The antitrust economics of payment card interchange fees". *Antitrust LJ* 73:571.

Levin, Dan, and James L Smith. 1996. "Optimal reservation prices in auctions". *The Economic Journal* 106 (438): 1271–1283.

Li, Xiaohu. 2005. "A note on expected rent in auction theory". *Operations Research Letters* 33 (5): 531–534.

Lu, Jay. 2019. "Bayesian identification: a theory for state-dependent utilities". *American Economic Review* 109 (9): 3192–3228.

Marra, Marleen. 2019. *Pricing and fees in auction platforms with two-sided entry.* Tech. rep. Sciences Po Departement of Economics.

Mela, Carl F, Jason MT Roos, and Tulio Sousa. 2023. "Advertiser Learning in Direct Advertising Markets". *arXiv preprint arXiv:2307.07015.*

Myerson, Roger B. 1981. "Optimal auction design". *Mathematics of operations research* 6 (1): 58–73.

Ostrovsky, Michael, and Michael Schwarz. 2016. *Reserve prices in Internet advertising auctions: a field experiment. Typescript.*

Platt, Brennan C. 2017. "Inferring ascending auction participation from observed bidders". *International Journal of Industrial Organization* 54:65–88.

Quint, Daniel. 2017. "Common values and low reserve prices". *The Journal of Industrial Economics* 65 (2): 363–396.

Resnick, Paul, and Richard Zeckhauser. 2002. "Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system". In *The Economics of the Internet and E-commerce.* Emerald Group Publishing Limited.

Rochet, Jean-Charles, and Jean Tirole. 2003. "Platform competition in two-sided markets". *Journal of the european economic association* 1 (4): 990–1029.

Rothschild, Michael. 1974. "A two-armed bandit theory of market pricing". *Journal of Economic Theory* 9 (2): 185–202.

Simonsohn, Uri. 2010. “eBay’s crowded evenings: Competition neglect in market entry decisions”. *Management science* 56 (7): 1060–1073.

Strulov-Shlain, Avner. 2021. “More than a Penny’s Worth: Left-Digit Bias and Firm Pricing”. *Chicago Booth Research Paper* 19–22.

Tadelis, Steven, et al. 2023. *Learning, Sophistication, and the Returns to Advertising: Implications for Differences in Firm Performance.* Tech. rep. National Bureau of Economic Research.

Vickrey, William. 1961. “Counterspeculation, auctions, and competitive sealed tenders”. *The Journal of finance* 16 (1): 8–37.

Wang, Zhide, and Nathan Yang. 2024. “Identification of Structural Learning Models”. *Available at SSRN.*

Wang, Zhide, et al. 2024. “Retail Investment under Aggregate Fluctuations”. *Available at SSRN.*

Xu, Boya, Yiting Deng, and Carl Mela. 2022. “A scalable recommendation engine for new users and items”. *arXiv preprint arXiv:2209.06128.*

A Appendix

A1 Summary statistics and additional descriptive evidence

Table A1.1 shows selected summary statistics for the sample used in the data. To limit the effect of prediction error in estimated item values (described in more detail in 4.1), I drop all items with a standardized reserve price and standardized revenue greater than the 99th quantile of the respective variables.

Table A1.1: Summary statistics

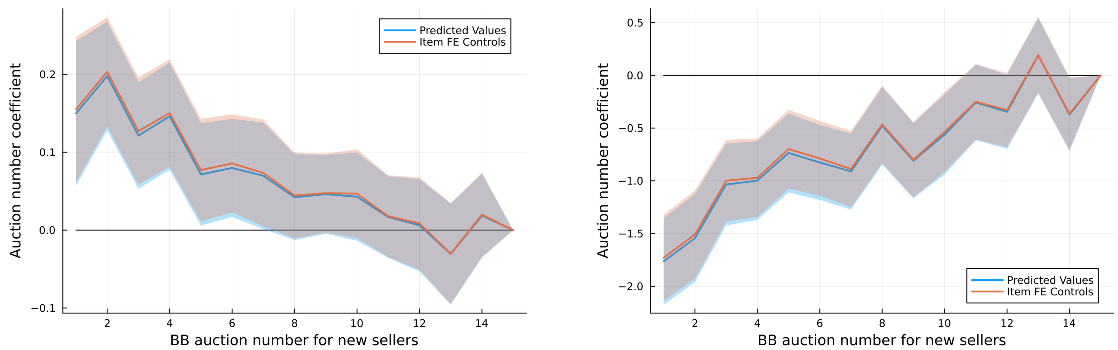
Variable	Mean	Std. Dev.	Minimum	Maximum
Minimum Bid	13.85	48.43	0.01	10,000
Reserve Price	15.39	139.4	0	68,000
Revenue	15.66	83.5	0	68,000
# Bidders	2.62	2.93	0	36
Sell	0.55	0.5	0	1
Fees	1.05	1.97	0	867.12

Notes: These statistics are from the 1,038,383 items included in the analysis data. The top 1% of the items (by standardized effective reserve price) have been removed from the analysis data.

Figure A1.1 estimates the same equation as Figure 2, but the dependent variables are the inverse hyperbolic sine (IHS) of effective reserve price and net revenue. I also restrict the sample to items where the seller has listed at least one other item with the same description, and run the regression with predicted item values and item fixed effects to compare the resulting estimates. The trends are quite similar whether using predicted item values or fixed effects, which suggests the predicted item values capture economically meaningful information. They are also similar to the trends in Figure 2, though with the caveat that the greater magnitude of the coefficients in panel (b) may be in part driven by re-listed items that were not sold the first time.[10]

[10] One limitation of the dataset is that I cannot observe seller inventories, so I cannot see how many of the identical items are true duplicates as opposed to relisting.

Figure A1.1: Regression coefficients α_k of auction experience on variables of interest

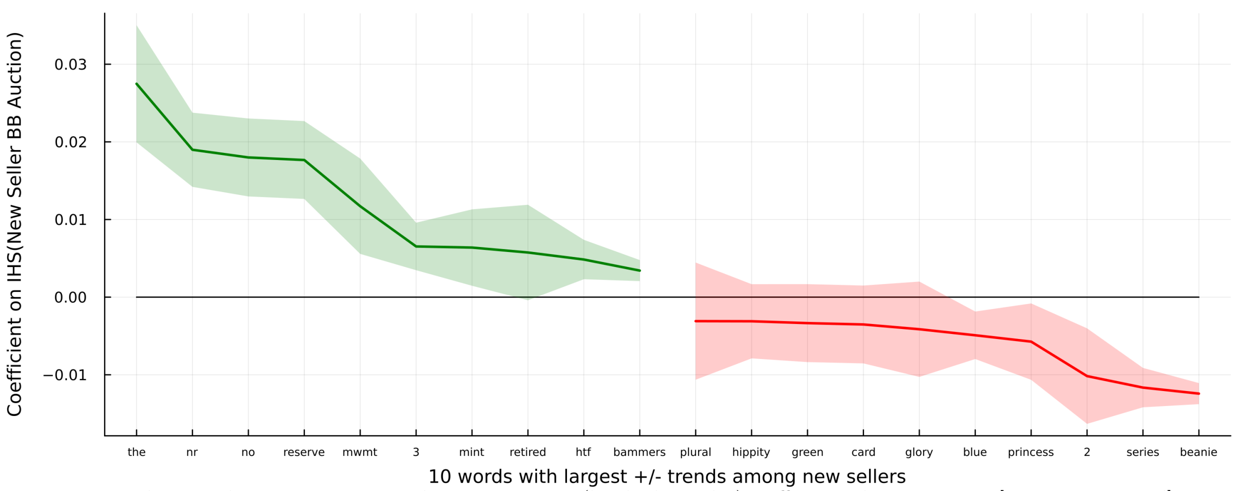


(a) IHS(Effective reserve) (b) IHS(Net revenue)

Notes: These regressions pool 1,639 new sellers' first 15 auctions with all auctions by 5,165 experienced sellers (defined as those with ≥47 auctions at the start of the data, which is the 75th percentile of initial experience). The sample is limited to sellers with at least 15 auctions in the data. The results are similar when using different values of T_{New}.

Figure A1.2 shows the words that most increased and decreased in their usage by new sellers in their first 15 auctions. While the trends are largely small, the words that most increased in frequency include "nr" (short for "no reserve") and "no" "reserve". This is consistent with sellers becoming more aware of the possible effect of their pricing decisions on bidder entry.

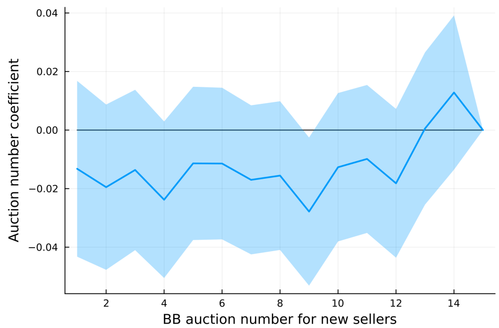
Figure A1.2: Trends in the frequency of words in item descriptions



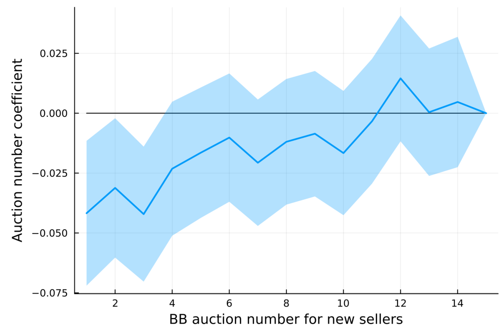
Notes: These are the 10 most positive and 10 most negative (by absolute value) coefficients when regressing $\mathbb{1}$[item contains word] on the inverse hyperbolic sine (IHS) of new sellers' auction number (among the first 15 auctions of sellers who have at least 15 auctions in the data or sellers with >75th percentile of experience at the start of the data), along with predicted item value, IHS(feedback count), feedback percentage, and seller and month fixed effects.

Figure A1.3 shows additional trends in non-price variables among sellers with at least 15 auctions. New sellers show some trends in the timing of an auction (panels (b) and (c)), where they favor shorter auctions that end on weekends. As shown in panel (d), new sellers also become less likely to feature items. To ensure that estimation is computationally tractable, and since these trends are generally smaller in magnitude relative to the baseline averages of each variable, I focus attention on the choice of reserve prices.

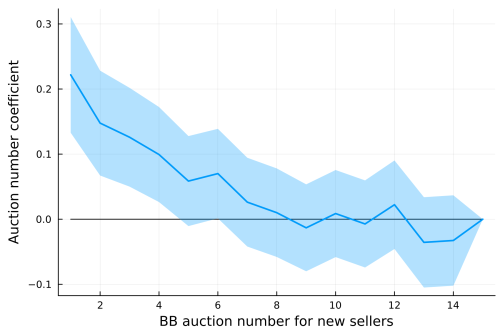
Figure A1.3: Trends in non-price variables among new sellers



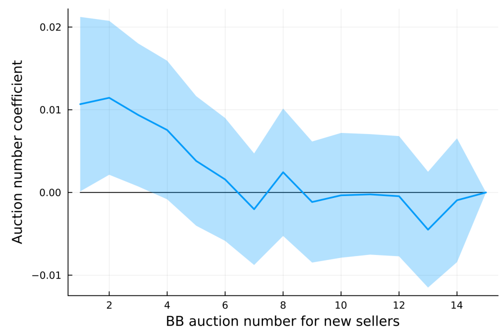
(a) End time in evening



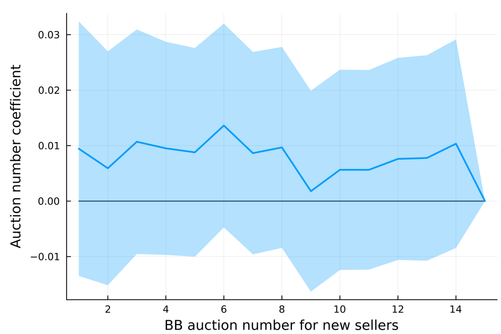
(b) End time on weekend



(c) Auction length (days)



(d) Featured item

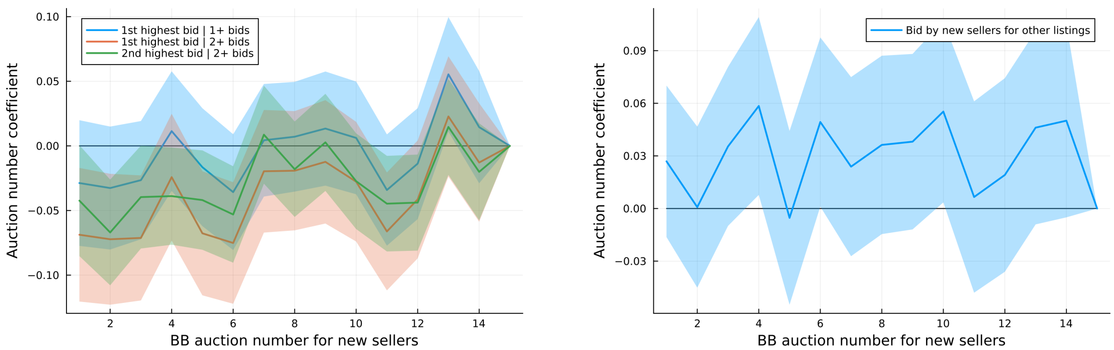


(e) Has secret reserve

Notes: These figures display coefficients of the first 15 auctions of new sellers in a regression of the various outcomes on predicted item values, seller feedback scores, inverse hyperbolic sine (IHS) of auction experience, and seller and month fixed effects. The regression is on new sellers with at least 15 auctions in the data and experienced sellers (defined as sellers with >75th percentile of experience at the start of the data).

Figure A1.4 estimates similar regressions for bids within auctions, though with the number of observed bids in each auction additional control. Panel (a) shows the trends in first and second highest bids faced by new sellers, since these are the only bids known to reflect the first and second highest values in an ascending IPV setting. Panel (b) shows the trend in bids placed by new sellers on other listings on or before their kth listing. The trend lines for both figures are relatively noisy and generally flat, though there is a slight upward trend in the bid values for auctions with 2 or more bidders in panel (a).

Figure A1.4: Regression coefficients α_k of auction experience on bids



(a) Trends in bids faced by new sellers

(b) Trends in bids placed by new sellers

Notes: These regressions pool bids from 1,639 new sellers' first 15 auctions with bids from all auctions by 5,165 experienced sellers (defined as those with $\geq$47 auctions at the start of the data, which is the 75th percentile of initial experience). The sample is limited to sellers with at least 15 auctions in the data. Panel (a) restricts the sample to all auctions with at least 1 or 2 bids (as specified in the legend), and panel (b) examines bids placed by new sellers before they list their kth item. Both panels control for seller fixed effects as in section 2.2, as well as fixed effects for the number of other observed bids in each auction.

A2 Proofs and derivations

Proof of Proposition 1

This closely follows the corollary in Marra (2019), though with the caveat that r is observed in this setting and thus directly impacts the Poisson mean Λ. First, denote

$$\pi_B(r \mid n, c) = \frac{1}{n} \cdot \mathbb{E}\Big[v_{n:n} - (1 + c_B^P) \max\{v_{(n-1):n}, r^*\} \;\Big|\; v_{n:n} \geq r^B\Big] \cdot (1 - F_B(r^B)^n)$$

Since F_B satisfies the strict monotone hazard rate property, Li (2005) implies that $\mathbb{E}[v_{(n+1):(n+1)} - v_{n:(n+1)}] < \mathbb{E}[v_{n:n} - v_{(n-1):n}]$; this holds when conditioning on r since r is set before the auction and does not vary with the number of bidders n that arrive. Additionally, $\frac{1}{n}(1 - F_B(r)^n) \geq$

$\frac{1}{n+1}(1-F_B(r)^{n+1})$,[11] so $\pi_B(r \mid n)$ is decreasing in n.

We temporarily abuse notation to write the probability mass at n given Λ as $p_n(\Lambda)$. Since arrival is Poisson, $p_n(\Lambda')$ first-order stochastically dominates $p_n(\Lambda)$ for $\Lambda' > \Lambda$. Increasing Λ therefore decreases $\sum_{n=1}^{\mathbf{N}_B-1} \pi_B(r \mid n,c)p_n(\Lambda)$ since $\pi_B(r \mid n)$ is monotonically decreasing in n. Thus, there exists a unique Λ that solves the zero profit condition.

Similarly, $\pi_B(r \mid n,c)$ is decreasing in r, since the measure of the set $\{v_{n:n} \geq r^B\}$, the probability of winning $1-F_B(r^B)^n$, and the winner's expected surplus are all all decreasing in r. Since higher r corresponds to strictly lower surplus, Λ is strictly decreasing in r.

Proof of Proposition 3

We differentiate the first-order condition with respect to v_0:

$$\frac{\partial^2 \Pi(v_0, r \mid b, c)}{\partial r \partial v_0} = (1-c_S^P)R_{rr}(r^*(v_0))\frac{\partial r^*(v_0)}{\partial v_0} + K_r(r^*(v_0) \mid b) + v_0 K_{rr}(r^*(v_0))\frac{\partial r^*(v_0)}{\partial v_0}$$

Rearranging, we have

$$\frac{\partial r^*(v_0)}{\partial v_0} = \frac{K_r(r^*(v_0) \mid b)}{-[(1-c_S^P)R_{rr}(r^*(v_0)) + v_0 K_{rr}(r^*(v_0))]}$$

The denominator is positive at the interior optimum because it is the negative second-order condition of the seller's profit maximization problem. Thus, $\frac{\partial r^*}{\partial v_0}$ when $K_r(\cdot \mid b) > 0$. Further, since the inverse of an increasing function is itself increasing, the virtual type function $\psi(\cdot \mid b, c)$ is monotonic increasing.

Lemma 1. Let $p_n(x \mid \delta) = \frac{1}{n!}\exp(-\Lambda(x \mid \delta))\Lambda(x \mid \delta)^n$, where $\Lambda(x \mid \delta) = \exp(\delta_1 + \delta_2 \rho(x))$ and and $\rho(x^*) = 0$. Then for each $k = 1, 2, ...,$ $\frac{\partial^k}{\partial x^k} p_n(x^* \mid \delta)$ has the form

$$\sum_{\ell=0}^{k} \delta_2^\ell \cdot h_{\ell,k}\left(n, \delta_1, \left\{\frac{\partial^t}{\partial x^t}\rho(x^*)\right\}_{t=1}^{k-\ell+1}\right)$$

[11] To see this, first note that $F_B(r) \in [0,1]$, and the expression is equivalent to showing $(1+\frac{1}{n})\frac{1-F_B(r)^n}{1-F_B(r)^{n+1}} \geq 1$. Let $x \in [0,1]$, and note that $1-x^n = (1-x)b_n$, where $b_n \equiv \sum_{k=0}^{n-1} x^k$. Then $\frac{1-x^n}{1-x^{n+1}} = \frac{b_n}{b_n+x^n} = \frac{1}{1+\frac{x^n}{b_n}}$, and $\frac{1}{n} \geq \frac{x^n}{b_n}$, since $\sum_{k=0}^{n-1} x^k \geq \sum_{k=0}^{n-1} x^n$.

where each $h_{\ell,k}$ is known.

Proof

We first show that for every $k = 1, 2, ...,$ $\frac{\partial^k}{\partial x^k} p_n(x \mid \delta)$ has the form

$$\sum_{\ell=0}^{k} \delta_2^{\ell} \cdot \tilde{h}_{\ell,k}\left(n, \Lambda(x \mid \delta), \left\{\frac{\partial^t}{\partial x^t}\rho(x)\right\}_{t=0}^{k-\ell+1}\right)$$

for known $\tilde{h}_{\ell,k}$. Beginning with $k = 1$, note that

$$\begin{aligned}\frac{\partial^1}{\partial x^1} p_n(x \mid \delta) &= \delta_2 \cdot p_n(x \mid \delta) \cdot [n - \Lambda(x \mid \delta)] \cdot \rho'(x) \\ &\equiv \delta_2 \cdot \tilde{h}_{1,1}(n, \Lambda(x \mid \delta), \rho'(x))\end{aligned}$$

since p_n is a known function of Λ. Now suppose $\frac{\partial^k}{\partial x^k} p_n(x \mid \delta)$ has the form above. Then using the shorthand $\tilde{h}_{j,\ell,k}$ to denote the first derivative with respect to the jth argument of $\tilde{h}_{\ell,k}$, note that by the chain rule,

$$\begin{aligned}&\frac{\partial}{\partial x}\sum_{\ell=0}^{k} \delta_2^{\ell} \cdot \tilde{h}_{\ell,k}\left(n, \Lambda(x \mid \delta), \left\{\frac{\partial^t}{\partial x^t}\rho(x)\right\}_{t=0}^{k-\ell+1}\right) \\ = \quad & \sum_{\ell=0}^{k} \delta_2^{\ell} \cdot \left[\tilde{h}_{2,\ell,k}\left(n, \Lambda(x \mid \delta), \left\{\frac{\partial^t}{\partial x^t}\rho(x)\right\}_{t=0}^{k-\ell+1}\right) \cdot \Lambda(x \mid \delta) \cdot \rho'(x) \cdot \delta_2 \right. \\ & \left. + \tilde{h}_{3,\ell,k}\left(n, \Lambda(x \mid \delta), \left\{\frac{\partial^t}{\partial x^t}\rho(x)\right\}_{t=0}^{k-\ell+1}\right)^{\intercal} \left\{\frac{\partial^{t+1}}{\partial x^{t+1}}\rho(x)\right\}_{t=1}^{k-\ell+2}\right] \\ \equiv & \sum_{\ell=0}^{k+1} \delta_2^{\ell} \cdot \tilde{h}_{\ell,k+1}\left(n, \Lambda(x \mid \delta), \left\{\frac{\partial^t}{\partial x^t}\rho(x)\right\}_{t=0}^{k-\ell+1}\right)\end{aligned}$$

Evaluating the expression above at x^*, where $\rho(x^* = 0)$ and therefore $\Lambda(x^* \mid \delta) = \exp(\delta_1)$, yields the desired result.

Proof of Proposition 5

By assumption (i), restricting attention to all sellers with the same history $\mathcal{H}$ is equivalent to restricting attention to sellers with identical beliefs b.

By Proposition 3, $\psi(\cdot \mid b, c)$ is increasing. Using the seller first-order condition, a change-of-variables can be applied to the reserve price distribution to write it in terms of the virtual type function and the known seller value distribution:

$$\mathbb{P}[r \le x] = \mathbb{P}[v_0 \le \psi(x \mid b, c)] = F_S(\psi(x \mid b, c))$$

Further, the selection rule for seller entry implies $\bar{v}(b,c) \ge v_0$. Taken together, the probability that a reserve price is less than or equal to x, conditional on beliefs b and cost vector c, is $F_S(\psi(x \mid b,c))/F_S(\bar{v}(b,c))$. Since by assumption (ii) the entry threshold $\bar{v}(b,c)$ is known, inverting the empirical reserve price distribution $\phi(x \mid \mathcal{H})$ of sellers with history $\mathcal{H}$ yields the virtual type function:

$$\psi(x \mid b, c) = F_S^{-1}\Big[\phi(x \mid \mathcal{H}) \cdot F_S(\bar{v}(b,c))\Big]$$

for all x such that $\psi(x \mid b, c) < \bar{v}(b,c)$. By assumption (iv) this support includes a positive-measure interval including some value x^* for which $\rho(x^*) = 0$; in what follows we restrict attention to this interval.

The virtual type function $\psi(x \mid b, c)$ is proportional to the ratio of $R_r(x \mid b)$ and $K_r(x \mid b)$, and its derivatives are

$$\frac{\partial^k}{\partial x^k}\psi(x \mid b, c) = -(1 - c_S^P)\frac{\partial^k}{\partial x^k}\frac{R_r(x \mid b)}{K_r(x \mid b)} = -(1 - c_S^P)\sum_{\ell=0}^{k}\binom{k}{\ell}\Big(\frac{\partial^{k-\ell}}{\partial x^{k-\ell}}R_r(x \mid b)\Big)\cdot\Big(\frac{\partial^\ell}{\partial x^\ell}\big(K_r(x \mid b)\big)^{-1}\Big)$$

where by Faà di Bruno's formula

$$\frac{\partial^\ell}{\partial x^\ell}\big(K_r(x \mid b)\big)^{-1} = \sum_{t=0}^{\ell}\frac{(-1)^t \cdot t!}{(K_r(x \mid b)^{t+1})}B_{\ell,t}\Big(\frac{\partial}{\partial x}K_r(x \mid b),\ \ldots,\ \frac{\partial^{\ell-t+1}}{\partial x^{\ell-t+1}}K_r(x \mid b)\Big)$$

in which $B_{\ell,t}$ are Bell polynomials. In turn, for both functions $G \in \{R, K\}$, we have

$$\frac{\partial^k}{\partial x^k}G(x \mid b) = \sum_{\ell=0}^{k}\binom{k}{\ell}\Bigg(\sum_{n=0}^{\infty}\Big(\frac{\partial^{k-\ell}}{\partial x^{k-\ell}}G_n(x)\Big)\cdot\Big(\frac{\partial^\ell}{\partial x^\ell}p_n(x \mid b)\Big)\Bigg)$$

Thus, the kth derivative of the virtual type function is a known function of the 0,...,kth derivatives of R_n, K_n, and p_n.

We now turn our attention to the Poisson mass function and its derivatives, which are the only arguments of ξ_k that depend on beliefs b. Expanding $\frac{\partial^k}{\partial x^k} p_n(x \mid b)$ and imposing independence between the marginal beliefs about the two parameters yields

$$\frac{\partial^k}{\partial x^k} p_n(x \mid b) = \int \int \left[\frac{\partial^k}{\partial x^k} p_n(x \mid \delta) \right] b_{\delta_1}(\delta_1) b_{\delta_2}(\delta_2) d\delta_1 d\delta_2$$

Evaluating this at x^* and applying Lemma 1, this can be expanded to yield

$$\sum_{\ell=0}^{k} \mathbb{E}_{b_{\delta_2}}[\delta_2^\ell] \cdot \hat{h}_{\ell,k}(n)$$

where $\hat{h}_{\ell,k}(n) \equiv \int h_{\ell,k}(n, \delta_1, \{\frac{\partial^t}{\partial x^t}\rho(x^*)\}_{t=1}^{k-\ell+1}) b_{\delta_1}(\delta_1) d\delta_1$ is known under the assumption that b_{δ_1} is known. Note that $\mathbb{E}_{b_{\delta_2}}[\delta_2^\ell]$ does not depend on n, so this term can be pulled out of all the infinite sums in which it appears, i.e. for $G \in \{R, K\}$

$$\begin{aligned}
\frac{\partial^k}{\partial x^k} G(x^* \mid b) &= \sum_{\ell=0}^{k} \binom{k}{\ell} \left(\sum_{n=0}^{\infty} \Big(\frac{\partial^{k-\ell}}{\partial x^{k-\ell}} G_n(x^*) \Big) \cdot \Big(\sum_{t=0}^{\ell} \mathbb{E}_{b_{\delta_2}}[\delta_2^t] \cdot \hat{h}_{t,\ell}(n) \Big) \right) \\
&= \mathbb{E}_{b_{\delta_2}}[\delta_2^k] \left(\sum_{n=0}^{\infty} G_n(x^*) \hat{h}_{k,k}(n) \right) + \sum_{\ell=0}^{k-1} \binom{k}{\ell} \left(\sum_{n=0}^{\infty} \Big(\frac{\partial^{k-\ell}}{\partial x^{k-\ell}} G_n(x^*) \Big) \cdot \Big(\sum_{t=0}^{\ell} \mathbb{E}_{b_{\delta_2}}[\delta_2^t] \cdot \hat{h}_{t,\ell}(n) \Big) \right)
\end{aligned}$$

Note the $k-1$th derivative of the virtual type function is a function of the kth raw moment of b_{δ_2}. By assumption, the $k-1$th derivative of $\psi(x^* \mid b, c)$ is invertible in the coefficient of this first term, yielding identification of the kth raw moment from the $k-1$th derivative of $\psi(x^* \mid b, c)$ and knowledge of lower-order moments. Since b_{δ_2} satisfies the Carleman condition, its moments uniquely characterize the distribution, and b_{δ_2} is identified up to the $\bar{k}$th moment.

A3 Demand estimation details

I use text data from item descriptions to estimate the average value for each item. Since item descriptions are seller-provided, there is significant variation in how words are spelled, which poses a challenge for tractably estimating item values. To address this, I manually created a crosswalk of individual words to their apparent intended word to decrease the dimensionality of the space item descriptions (e.g., replacing “beaneis” and “babys” with “beanies” and “babies”). I then constructed a dictionary of the 5,368 words that appear at least 10 times in the cleaned item descriptions. I also include indicators for each month in the dataset.

I tested multiple neural network architectures for γ via out-of-sample validation and with built-in dropout layers to find the architecture that achieved the lowest out-of-sample loss using the likelihood derived in A5. In particular, I use 80% of the sample to train the model, 10% of the sample to test out-of-sample loss during training, and the remaining 10% of the sample for out-of-sample validation after training. I used an early stopping rule to determine the number of epochs with which to train the full model: I use the smallest number of epochs after which the testing loss fails to improve for 10 consecutive epochs. I then select the architecture with the lowest validation loss. The resulting architecture has 9,111,233 parameters; additional information on the various architectures is presented in Table A3.1. This table shows that the nonparametric specifications outperform the parametric model in the first line of the table. Larger numbers of parameters generally improve validation loss, though there are diminishing returns to increased complexity. The empirical results of the paper are similar using different architectures.

Table A3.1: Neural network architectures and performance

Model	# 1st-Layer Nodes	# Parameters	Train Loss	Test Loss	Val Loss	Epochs	R^2
1	-	5,374	2.6013	2.6165	2.6104	46	0.691
2	512	2,900,321	1.0885	1.2647	1.2680	34	0.984
2	1,024	5,782,881	0.9965	1.2264	1.2374	38	0.992
2	1,536	8,665,441	1.0006	1.2316	1.2752	30	0.989
3	512	3,083,969	1.0509	1.2638	1.2791	47	0.987
3	1,024	6,097,601	0.9144	1.1934	1.2007	43	0.993
3	1,536	9,111,233	0.8590	1.1774	1.1933	46	-

Notes: Model 1 is fully parametric, with a Gaussian distribution for log-values and no hidden nodes (i.e., log-values are modeled as a linear combination of word-specific fixed effects). Models 2 and 3 both use 5-component Gaussian mixture models, with a varying number of nodes in the first hidden layer; both have 4 hidden layers with corresponding dropout of 50%, 40%, 30%, and 20% for each. The second through fourth hidden layers contain 256, 64, and 16 nodes for model 2 and 512, 128, and 32 nodes for model 3. The number of parameters is the total number of trained parameters in each specification. The train, test, and validation loss columns denote the loss of each of the 80%, 10% and 10% samples used in comparing each of the models. The number of epochs is chosen via early stopping, since subsequent training after the listed number of epochs yields no test loss improvement for at least 10 epochs. The R^2 is taken from regressing the fitted item values from each architecture on the architecture with the lowest validation loss.

A4 Orthogonalization of the likelihood function

This section derives a method to estimate the true parameter ϑ_0 without bias due to estimation error for the nonparametric component γ_0. I denote the log-likelihood as ℓ; its derivation is shown in the following section. All relevant data for this demand-side likelihood is abbreviated as $\mathbf{D}_d$ to differentiate it from the data $\mathbf{D}$ that is used by sellers in updating their beliefs.

Denote the score function for the structural parameters ϑ_0 as

$$g(\vartheta \mid \gamma, \mathbf{D}_d) = \frac{\partial \ell(\vartheta, \gamma \mid \mathbf{D}_d)}{\partial \vartheta}$$

and note that $\mathbb{E}[g(\vartheta_0 \mid \gamma_0, \mathbf{D}_d)] = 0$. To derive a Neyman orthogonal score $g^*(\vartheta \mid \gamma, \mathbf{D}_d)$ for the average score $\mathbb{E}[g(\vartheta_0 \mid \gamma_0, \mathbf{D}_d)]$, Ichimura and Newey (2022) provide a method for finding a candidate first-stage influence function that will be added to the original score. I follow the steps in their Proposition 1 to show how this applies to a setting with both low and high dimensional parameters, where we orthogonalize with respect to the high dimensional parameter.

By way of notation, γ_0 as the true high-dimensional parameter under the true distribution function, and γ_τ is the perturbation in the direction of some alternative $\tilde{\gamma}$ (i.e., $\gamma_\tau = (1-\tau)\gamma_0 + \tau\tilde{\gamma}$). The Gateaux derivative $\frac{\partial}{\partial \tau}$ is the derivative with respect to τ from above evaluated at zero ($\tau \downarrow 0$). I assume that $\mathbb{E}[\frac{\partial}{\partial a}\frac{\partial}{\partial a}\ell(\vartheta, \gamma(X) + a \mid \mathbf{D}_d) \mid X = x] = 0$, which implies $\mathbb{E}[b(X)\ell(\vartheta, \gamma(X) + a \mid \mathbf{D}_d) \mid X = x] = 0$ for all b.

In this setting, Assumptions 1 and 2 of Ichimura and Newey (2022) are that there exist $\alpha_1(\vartheta \mid x)$ and $\alpha_2(\vartheta \mid x)$ with finite variance (and where α_2 is bounded away from zero) such that

$$\begin{aligned}
\frac{\partial}{\partial \tau}\mathbb{E}\left[\frac{\partial}{\partial \vartheta}\ell(\mathbf{D}_d, \vartheta, \gamma_\tau(X))\right] &= \frac{\partial}{\partial \tau}\mathbb{E}\left[\alpha_1(\vartheta \mid X)\gamma_\tau(X)\right] \\
\frac{\partial}{\partial \tau}\mathbb{E}\left[b(X)\frac{\partial}{\partial \gamma}\ell(\mathbf{D}_d, \vartheta, \gamma_\tau(X))\right] &= \frac{\partial}{\partial \tau}\mathbb{E}\left[b(X)\alpha_2(\vartheta \mid X)\gamma_\tau(X)\right]
\end{aligned}$$

By the chain rule and iterated expectations on the score above, we have

$$\begin{aligned}
\alpha_1(\vartheta \mid x) &= \mathbb{E}\left[\frac{\partial g(\vartheta \mid a, \mathbf{D})}{\partial a}\Big|_{a=\gamma(X)}\middle| X = x\right] \\
\alpha_2(\vartheta \mid x) &= \mathbb{E}\left[\frac{\partial^2 \ell(\vartheta \mid a, \mathbf{D})}{\partial a^2}\Big|_{a=\gamma(X)}\middle| X = x\right]
\end{aligned}$$

Writing the derivative of the likelihood with respect to the scalar output of γ as

$$\tilde{g}(\vartheta \mid \gamma, \mathbf{D}) = \frac{\partial \ell(\vartheta \mid a, \mathbf{D})}{\partial a}\Big|_{a=\gamma(X)}$$

we can combine these terms to form the orthogonal score

$$g^{*}(\vartheta \mid \gamma, \mathbf{D}_d) = g(\vartheta \mid \gamma, \mathbf{D}_d) - \alpha_1(\vartheta \mid x) \cdot \alpha_2(\vartheta \mid x)^{-1} \cdot \tilde{g}(\vartheta \mid \gamma, \mathbf{D}) \tag{12}$$

This orthogonal score may then be used to estimate θ while removing bias due to the plug-in estimator γ_0.

The nuisance parameters α_1 and α_2 are projections of second derivatives of ℓ onto the space of covariates X that enter γ. Unlike regression settings, each depends on the structural parameters ϑ; this is similar to Example 3 of Chernozhukov et al. (2022). As in that setting, initial estimators $\hat{\gamma}$ and $\hat{\theta}$ can be constructed using sample splitting, and then plugged into α_1 and α_2 to get predicted values and estimate the conditional expectations $\hat{\alpha}$ using nonparametric regression on X. These "plugin" estimators form the nuisance parameter $\hat{\alpha}(x) = \hat{\alpha}_1(\hat{\vartheta} \mid x) \cdot \hat{\alpha}_2(\hat{\vartheta} \mid x)^{-1}$ that yields a version of equation (12) that will be used in estimation (I omit the multiple indices used in sample splitting for ease of exposition):

$$g^{*}(\vartheta \mid \gamma, \mathbf{D}_d) = g(\vartheta \mid \gamma, \mathbf{D}_d) - \hat{\alpha}(x) \cdot \tilde{g}(\vartheta \mid \gamma, \mathbf{D}_d)$$

This orthogonal moment can be used as in standard GMM both to estimate ϑ without bias and construct the asymptotic variance matrix of the structural parameters. I follow the steps in Chernozhukov et al. (2022) with threefold sample splitting.

A5 Likelihood derivation: demand side

I now derive the demand side likelihood as presented in equation (10). Subscripts for j and t will be omitted where possible (since this focuses on the likelihood contribution for any given auction), as well as the dependence on parameters ϑ_d, to streamline notation.

The likelihood that no bids are observed is simply the likelihood that the highest bid (integrated over the arrival distribution) is smaller than the minimum bid. Thus, the likelihood contribution for observing no bids is

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{\Lambda^n e^{-\Lambda}}{n!} \underbrace{F_B(m)^n}_{\mathbb{P}[\text{all bids below } m]} &= \\
&= e^{-\Lambda[1-F_B(m)]} \sum_{n=0}^{\infty} \frac{[\Lambda F_B(m)]^n e^{-[\Lambda F_B(m)]}}{n!} \\
&= e^{-\Lambda[1-F_B(m)]}
\end{aligned}$$

where the third equality holds since the sum is the integral of a Poisson density with mean $\Lambda F_B(m)$.

The likelihood contribution from auctions with one observed bidder uses the fact that one bid is not censored, but the other $n-1$ are. For any n bidders that arrive,

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{\Lambda^n e^{-\Lambda}}{n!} \underbrace{n F_B(m)^{n-1} f_B(v^{(1)})}_{\mathbb{P}[\text{only 1 bid above } m]} &= \\
&= \frac{f_B(v^{(1)})}{F_B(m)} e^{-\Lambda[1-F_B(m)]} \sum_{n=0}^{\infty} \frac{[\Lambda F_B(m)]^n e^{-[\Lambda F_B(m)]}}{n!} n \\
&= f_B(v^{(1)}) \Lambda e^{-\Lambda[1-F_B(m)]}
\end{aligned}$$

where the second equality holds since the summand is 0 for $n = 0$, and the last equality holds since the sum is the first moment of the Poisson distribution with mean $\Lambda F_B(m)$.

The last case, where $N_j \geq 2$, combines all possible arrival orders with at least 2 bidders. The precise values of other bidders are not necessary to construct the partial likelihood; the two highest bids provide enough information about the arrival process.

$$
\begin{aligned}
\sum_{n=2}^{\infty} \frac{\Lambda^n e^{-\Lambda}}{n!} & \underbrace{f_B(v^{(1)} \mid v^{(2)}) f_B(v^{(2)} \mid n \text{ bids})}_{\mathbb{P}[\text{2 highest bids}]} = \\
&= f_B(v^{(1)} \mid v^{(2)}) \sum_{n=2}^{\infty} \frac{\Lambda^n e^{-\Lambda}}{n!} n(n-1)(1 - F_B(v^{(2)})) F_B(v^{(2)})^{n-2} f_B(v^{(2)}) \\
&= f_B(v^{(1)} \mid v^{(2)}) \frac{(1 - F_B(v^{(2)}))}{F_B(v^{(2)})} \frac{f_B(v^{(2)})}{F_B(v^{(2)})} e^{-\Lambda[1 - F_B(v^{(2)})]} \\
&\qquad \cdot \sum_{n=0}^{\infty} \frac{[\Lambda F_B(v^{(2)})]^n e^{-[\Lambda F_B(v^{(2)})]}}{n!} (n^2 - n) \\
&= f_B(v^{(1)}) f_B(v^{(2)}) \Lambda^2 e^{-\Lambda[1 - F_B(v^{(2)})]}
\end{aligned}
$$

where the third equality follows since $n^2 - n = 0$ for $n = 0, 1$, the fourth equality comes from the difference of the first and second raw moments of the Poisson distribution and further simplification. Combining the likelihood component of each case ($N = 0$, $N = 1$, or $N \geq 2$) with the density of the reserve price, we obtain equation (10).

Test with simulated data

For the simulations, I use a modified version of the data-generating process various architectures for the item value index γ. In the first architecture, I assume item j's log-value γ_j is known (i.e. $\gamma(\gamma_j) = \gamma_j$). In the following, I assume item values are a function of 50 indicator variables, each of which is randomly generated with average probability 0.1. I allow item values to be generated from a dense neural network (mapping from 50 indicator variables to two hidden layers of 10 nodes each before outputting to a scalar).

The number of bidders is Poisson distributed with mean $\Lambda_j = \exp(\delta_{0,1} + \delta_{0,2}\gamma_j)$, so in these simulations bidder arrival does not depend on the reserve price. I also set $m = r$ in the simulations, so the minimum bid and reserve price are the same. I model $r_j \sim \mathcal{N}(\mu_r, \sigma_r^2)$ and $v_{ij} \sim \mathcal{N}(0, \sigma_B^2)$. I report the results for this set of simulations in Table A5.1.

Table A5.1: Simulations for maximum likelihood estimation (demand)

	Regression: γ_0 on $\hat{\gamma}$		Structural Parameters				
	Intercept	Coef	σ_B^2	μ_r	σ_r^2	$\delta_{0,1}$	$\delta_{0,2}$
True values	0.0	1.0	0.5	0.25	0.5	1.0	0.5
Known γ_0							
N = 2,000	-	-	0.498	0.251	0.499	1.019	0.469
	-	-	(0.009)	(0.012)	(0.008)	(0.031)	(0.047)
N = 10,000	-	-	0.5	0.25	0.5	1.022	0.473
	-	-	(0.004)	(0.005)	(0.004)	(0.012)	(0.02)
Estimated γ_0 (uncorrected)							
N = 2,000	0.002	0.976	0.501	0.261	0.491	1.095	0.748
	(0.112)	(0.091)	(0.06)	(0.112)	(0.02)	(0.155)	(0.164)
N = 10,000	0.031	1.099	0.55	0.283	0.528	1.049	0.749
	(0.055)	(0.04)	(0.05)	(0.087)	(0.014)	(0.088)	(0.108)
Estimated γ_0 (orthogonalized)							
N = 2,000	-0.022	0.91	0.661	0.235	0.625	0.809	0.163
	(0.134)	(0.119)	(0.098)	(0.136)	(0.042)	(0.243)	(0.296)
N = 10,000	0.035	1.069	0.59	0.273	0.556	0.985	0.554
	(0.07)	(0.054)	(0.062)	(0.101)	(0.021)	(0.105)	(0.153)

Notes: Average (standard deviation) parameters are from 100 simulations for each case. Starting values were chosen randomly using Julia's Flux package initializations.

A6 Likelihood approach: supply side belief estimation

Several functions (e.g. R, K, and their derivatives with respect to r, and expectations with respect to belief densities or bidder values) involve multiple integrals and/or summations, and are therefore infeasible to compute repeatedly for all possible parameters ϑ_s (the parameters of the seller value and cost distributions) and ϑ_b (the parameters of the seller prior beliefs). I use dense neural networks to approximate several functions used in the estimation procedure.

Each neural network maps from $\mathbb{R}^Q$ to $\mathbb{R}$, and each is composed of one input layer, 9 hidden layers, and one output layer. The activation function for each hidden layer is leakyrelu, and the number of nodes from input layer to output layer for each neural network is: Q, 50, 100, 100, 200, 300, 3000, 300, 200, 100, 100, 50, 1. The activation function for the output layer is listed with the associated function below.

To construct each approximation, I generate datasets on which to train each neural network for various parameter values. The bounds of each variable used in the approximations are chosen to cover the empirical support of the corresponding variables where they are observed (e.g., $\delta_{j,0,1}$)

and sufficiently large support where they are unobserved (e.g. seller prior parameters). I use 99% of each dataset for training and 1% for holdout validation. I train each neural network on the respective training datasets in batches of 50 for 50 epochs before training the network on the full training dataset for 25 epochs; I exit training early if the mean square error of the holdout sample is less than 1e-5. The approximations (in bold) are constructed in the following order, with additional details listed for each approximation and the construction of the associated datasets. Each function is fit by minimizing mean square prediction error, though in some cases transformations are applied to improve accuracy for some parameter values.

1. *Functions with δ known.*
   (a) *Evaluate revenue and keep probabilities.* Using the estimated bidder value distribution and arrival parameters, I first evaluate R_n and K_n for $n = 0, 1, ..., 150$. I then construct 316 Chebyshev nodes in each dimension for $r \in [0.01, 6.25]$ and $\Lambda \in [-5, \ln(150)]$ and evaluate R and K, respectively, by taking their dot product with $\{p_n(r, \Lambda)\}_{n=0}^{150}$ evaluated at each node (I chose 316 because Floor($100{,}000^{0.5}$) = 316).
      i. **Expected revenue** R (exponential activation). Inputs: r and $\delta_{j,0,1}$.
      ii. **Keep probability** K (sigmoid activation). Inputs: r and $\delta_{j,0,1}$.

   (b) *Search for optimal reserve price.* I construct 316 Chebyshev nodes in each dimension for $v_0 \in [-1.25, 6.25]$ and $\delta_{j,0,1} \in [-5, \ln(150)]$ and search for the optimal reserve price r^* in 0.01, 0.02, ..., 6.25 along with the expected profit and seller surplus (profit minus outside option value) at the optimum.
      i. **Virtual type** ψ (identity activation). Inputs: r^* and $\delta_{j,0,1}$.
      ii. **Optimal reserve price** ψ^{-1} (identity activation). Inputs: v_0 and $\delta_{j,0,1}$.
      iii. **Expected surplus** $\Pi^* - v_0$ (exponential activation) Inputs: v_0 and $\delta_{j,0,1}$. Since expected surplus is positive when entry costs are zero, I minimize the mean square prediction error of the *log* expected surplus. This increases the relative accuracy of predicted expected surplus where it is small, which is important for precisely approximating the entry threshold in the next step.

   (c) *Entry threshold.* I then construct 316 Chebyshev nodes in each dimension for $c_E \in [0, 0.5]$ and $\delta_{j,0,1} \in [-5, \ln(150)]$ and search for the maximum $v_0 \in [-1.25, 6.25]$ such that expected surplus is weakly positive. Since expected surplus is monotonic in v_0, I use a binary search algorithm (i.e., evaluating expected surplus at the midpoint of $[-1.25, 6.25]$, determining whether $\bar{v}$ lies above or below the midpoint, and iterating with additional intervals) until the difference in successive iterations is less than 0.01.

   i. **Entry threshold** $\bar{v}$ (identity activation). Inputs: c_E and $\delta_{j,0,1}$.

2. *Functions with δ unknown.* I approximate the following functions for the full model with 2-dimensional unknown parameter δ_0 and the arrival coefficient model where only $\delta_{0,2}$ is unknown. The prior parameters are the mean $\{\mu_{0,1}, \mu_{0,2}\}$, standard deviations $\sigma_{0,1}$ and $\sigma_{0,2}$, and correlation $\tilde{\rho}_0$. Due to the higher dimensionality due to the belief parameters, I sample 100,000 input values for each step rather than constructing a grid of Chebyshev nodes. Unless otherwise specified, each input is sampled uniformly on the stated support.

   (a) *Search for optimal reserve price.* I sample $v_0 \sim F_S$ (bounded on $[1.25, 6.25]$), $\delta_{j,0,1} \sim U[-1.5, \ln(150)]$, $\mu_{0,2} \sim U[-0.75, 0.75]$, $\sigma_{0,i} \sim U[0.01, \sqrt{0.5}]$ for $i = 1, 2$, and $\tilde{\rho}_0 \sim U[-0.95, 0.95]$ ($\sigma_{0,2}$ and $\tilde{\rho}_0$ are only sampled for the full model). I then search for the optimal reserve price r^* in 0.01, 0.02, ..., 6.25 along with the expected profit and seller surplus (profit minus outside option value) at the optimum, integrating over sellers' belief distributions to do so using Gauss-Hermite quadrature with 5 points in each dimension.[12]

      i. **Virtual type** ψ (identity activation). Inputs: r^*, $\delta_{j,0,1}$, $\mu_{0,2}$, $\sigma_{0,1}$, $\sigma_{0,2}$, $\tilde{\rho}_0$.

      ii. **Optimal reserve price** ψ^{-1} (identity activation). Inputs: v_0, $\delta_{j,0,1}$, $\mu_{0,2}$, $\sigma_{0,1}$, $\sigma_{0,2}$, $\tilde{\rho}_0$.

      iii. **Expected surplus** $\Pi - v_0$ (exponential activation). Inputs: v_0, $\delta_{j,0,1}$, $\mu_{0,2}$, $\sigma_{0,1}$, $\sigma_{0,2}$, $\tilde{\rho}_0$. As with step 2(c), I minimize mean square prediction error using the log expected surplus.

   (b) *Entry threshold.* I sample $c_E \sim U[0, 1]$ (bounded on $[1.25, 6.25]$), $\delta_{j,0,1} \sim U[-1.5, \ln(150)]$, $\mu_{0,2} \sim U[-0.75, 0.75]$, $\sigma_{0,i} \sim U[0.01, \sqrt{0.5}]$ for $i = 1, 2$, and $\tilde{\rho}_0 \sim U[-0.95, 0.95]$ ($\sigma_{0,2}$ and $\tilde{\rho}_0$ are only sampled for the full model). I use the same binary search algorithm as in 1(c) to find $\bar{v}$.

      i. **Entry threshold** $\bar{v}$ (identity activation). Inputs: c_E, $\delta_{j,0,1}$, $\mu_{0,2}$, $\sigma_{0,1}$, $\sigma_{0,2}$, $\tilde{\rho}_0$.

[12] This procedure yields candidate quadrature nodes $\ln(\Lambda)$ at which I evaluate the expected revenue and keep probaiblities. For $\ln(\Lambda) \in [-5, \ln(150)]$ I use the evaluated R_n and K_n as in step 1(a) above, and for $\ln(\Lambda) \notin [-5, \ln(150)]$ I extrapolate using the fitted values of R and K.

(c) *Updating process.* I sample 200,000 draws of $v_0 \sim F_S$ (bounded on $[1.25, 6.25]$), $Z'\lambda \sim U[-1.5, \ln(150)]$, $\mu_{0,i} \sim U[-0.75, 0.75]$ for $i = 1,2$, $\sigma_{0,i} \sim U[0.01, \sqrt{0.5}]$ for $i = 1,2$, and $\tilde{\rho}_0 \sim U[-0.95, 0.95]$ ($\sigma_{0,2}$ and $\tilde{\rho}_0$ are only sampled for the full model). I keep 100,000 draws for which $\delta_{j,0,1} = Z'\lambda + \mu_{0,1} \in [-5, \ln(150)]$. I then use the optimal reserve price function to construct r^* and and the transformed reserve price $\rho(r^*)$. I also simulate profit signals ϵ as in equation (6) from $\mathcal{N}(0, \sigma_\Pi^2)$, where σ_Π^2 is estimated using the empirical difference between profit signals and expected profit among experienced sellers.

I compute the updated beliefs via a Laplace approximation to the posterior for each observation. I first use a resilient backpropogation algorithm to search for the new maximum a posteriori estimates $\delta^*_{j,0,1}$ and $\mu^*_{0,2}$ given prior parameters, r^*, and ϵ. I then evaluate the updated covariance parameters ($\sigma^*_{0,1}$, $\sigma^*_{0,2}$, and $\tilde{\rho}^*_0$) by evaluating the Hessian of the posterior evaluated at the maximum a posteriori estimate. I drop all evaluations for which this algorithm returns either posterior parameters outside the simulation bounds or an invalid covariance matrix, and evaluate the neural networks using the resulting posterior parameters.

i. **Updating means $\mu_{0,1}$ and $\mu_{0,2}$** (identity activation). Inputs: ϵ, $\delta_{j,0,1}$, $\rho(r^*)$, $\mu_{0,2} \cdot \rho(r^*)$, $\sigma_{0,1}$, $\sigma_{0,2}$, $\tilde{\rho}_0$. Instead of approximating each updated parameter directly, I minimize mean square prediction error for the standardized difference $(\mu^*_{0,i} - \mu_{0,i})/\sigma_{0,i}$ for $i = 1,2$; this ensures more accurate update steps for the sellers' mean parameters when beliefs are more highly concentrated.

ii. **Updating standard deviations $\sigma_{0,1}$ and $\sigma_{0,2}$** (identity activation). Inputs: ϵ, $\delta_{j,0,1}$, $\rho(r^*)$, $\mu_{0,2} \cdot \rho(r^*)$, $\sigma_{0,1}$, $\sigma_{0,2}$, $\tilde{\rho}_0$. Instead of approximating each updated parameter directly, I minimize mean square prediction error for the ratio $(\sigma^*_{0,1})/\sigma_{0,i}$ for $i = 1,2$; this ensures more accurate update steps for the sellers' covariance parameters when beliefs are more highly concentrated.

iii. **Updating posterior correlation $\tilde{\rho}$** (identity activation). Inputs: ϵ, $\delta_{j,0,1}$, $\rho(r^*)$, $\mu_{0,2} \cdot \rho(r^*)$, $\sigma_{0,1}$, $\sigma_{0,2}$, $\tilde{\rho}_0$. Instead of approximating the updated parameter directly, I minimize mean square prediction error for the difference $\tilde{\rho}^*_0 - \tilde{\rho}_0$ for $i = 1,2$.

I chose both $\rho(r^*)$ and $\mu_{0,2} \cdot \rho(r^*)$ as inputs after experimenting with various architectures.